



**Titre:** Analyse de sensibilité dans un contexte de prévision du prix des métaux  
Title: métaux

**Auteur:** Abbassi Abbas Mahamat  
Author:

**Date:** 2011

**Type:** Mémoire ou thèse / Dissertation or Thesis

**Référence:** Abbas Mahamat, A. (2011). Analyse de sensibilité dans un contexte de prévision du prix des métaux [Mémoire de maîtrise, École Polytechnique de Montréal].  
Citation: PolyPublie. <https://publications.polymtl.ca/599/>

 **Document en libre accès dans PolyPublie**  
Open Access document in PolyPublie

**URL de PolyPublie:** <https://publications.polymtl.ca/599/>  
PolyPublie URL:

**Directeurs de recherche:** Bernard Clément, & Luc Adjengue  
Advisors:

**Programme:** Mathématiques appliquées  
Program:

UNIVERSITÉ DE MONTRÉAL

**Analyse de sensibilité dans un contexte  
de prévision du prix des métaux**

ABBASSI ABBASS MAHAMAT

DÉPARTEMENT DE MATHÉMATIQUES APPLIQUÉES ET GÉNIE INDUSTRIEL

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

MÉMOIRE PRÉSENTÉ EN VUE DE L'OBTENTION

DU DIPLÔME DE MAÎTRISE

(SCIENCES APPLIQUÉES)

MARS 2011

© ABBASSI ABBASS MAHAMAT, 2011.

UNIVERSITÉ DE MONTRÉAL

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

Ce mémoire intitulé:

Analyse de sensibilité comme outil de prévision et application au marché des métaux

Présenté par : ABBASS MAHAMAT, Abbassi, pour un mémoire

En vue de l'obtention du diplôme de : Maîtrise en Sciences Appliquées

A été dûment accepté par le jury d'examen constitué de :

M.LABIB Richard, Ph.D., président

M.CLEMENT Bernard, Ph.D., membre et directeur de recherche

M. ADJENGUE Luc, Ph.D., membre et codirecteur de recherche

M. Mario Lefebvre, Ph.D., membre

## DÉDICACE

*Je dédie ce mémoire à mes parents pour tous les sacrifices faits pour me permettre d'avoir l'éducation qu'ils n'ont jamais eue. Je salue leurs courages pour m'avoir permis de partir vers des contrées lointaines afin de pouvoir élever mon esprit. Je dédie également ce mémoire aux autres membres de ma famille qui dans leurs conseils et soutien j'ai toujours trouvé un puits pour me ressourcer, pour ces courts moments pleins de souvenirs qui m'ont permis d'être l'homme que je suis. A mes amis, camarades des nuits blanches passées à étudier, ils se reconnaîtront, merci pour le soutien et d'avoir à votre manière contribué en maintenant toujours la barre haute, au dépassement de mes limites; que cet esprit de compétition guide le reste de vos jours sur terre et fasse de vous les hommes que vous rêvez d'être.*

## **REMERCIEMENTS**

Je remercie mes directeurs de recherche, les professeurs Bernard Clément et Luc Adjengue, pour avoir cru en moi et m'avoir donné l'opportunité d'expérimenter cette facette de l'éducation qu'est la recherche. Je remercie également tous ceux qui de près ou de loin m'ont orienté dans cette recherche : mes collègues de la compagnie Bombardier pour avoir permis au novice que j'étais d'expérimenter mes capacités. Enfin, je remercie de tout mon être l'Éternel pour avoir été à mes côtés dans tous les moments et m'avoir guidé et honoré en me rappelant par ces mots simples selon lesquels ceux qui croient en lui ne seront jamais seuls.

## RÉSUMÉ

Ce mémoire explore le concept d'analyse de sensibilité dans un contexte de prévision. On s'intéresse à un indice du prix des métaux (codé sous le terme WPU10) publié par le bureau américain de la statistique du travail ou « Bureau of Labor Statistics » (BLS). Pour ce faire, on utilise la décomposition de l'indice WPU10 fournie par le BLS pour construire un indice synthétique répliquant son comportement que, dans une seconde phase nous utiliserons pour faire des prévisions à l'aide du principe de processus de diffusion développé en finance. On fait intervenir des notions telles que ceux de processus stochastiques, de factorisation de Cholesky et des techniques de simulations comme celle de Monte Carlo. Enfin, nous comparons le modèle développé en termes de flexibilité et de précision dans la prédiction à un modèle plus classique de prédiction de série temporelle.

## **ABSTRACT**

The main subject of this thesis is about the scientific principle of sensitivity analysis, where we studied the volatility of a metal price index published by the American Bureau of Labor Statistics (BLS) coded under the term WPU10. Thus we have developed a forecast model of the index of interest based on a Monte Carlo simulation using a multivariate random number matrix generated through a Cholesky factorization. In addition we have compared the efficiency of our forecast model to another model based on the study of time series using auto regression.

## TABLE DES MATIÈRES

DÉDICACE.....	III
REMERCIEMENTS .....	IV
RÉSUMÉ.....	V
ABSTRACT .....	VI
TABLE DES MATIÈRES .....	VII
LISTE DES TABLEAUX.....	X
LISTE DES FIGURES .....	XII
LISTE DES SIGLES ET ABRÉVIATIONS .....	XIII
LISTE DES ANNEXES .....	XIV
INTRODUCTION.....	1
CHAPITRE 1    REVUE DE LITTÉRATURE .....	3
1.1    Introduction .....	3
1.2    Méthodes de criblage .....	4
1.2.1    Méthode de criblage simple ou à très grande dimension .....	4
1.2.2    La méthode de Morris .....	6
1.3    Analyse de sensibilité locale .....	6
1.4    Analyse de sensibilité globale .....	7
CHAPITRE 2    MÉTHODES ET INDICES DE L'ANALYSE DE SENSIBILITÉ.....	9
2.1    Indices de sensibilité .....	9
2.1.1    Indices de sensibilité dans le cas du modèle linéaire .....	9
2.1.2    Indices de sensibilité cas du modèle non linéaire et monotone.....	14
2.1.3    Indices de sensibilité cas du modèle non linéaire et non monotone.....	15
2.2    Méthode d'analyse de sensibilité : forme ou fonction du modèle inconnue .....	21



2.2.1	Planification d'expérience.....	21
2.3	Conclusion.....	27
CHAPITRE 3 MODÉLISATION DU MARCHÉ DES MÉTAUX .....		29
3.1	Introduction .....	29
3.2	L'indice et l'institution.....	29
3.2.1	Présentation du BLS.....	29
3.2.2	Introduction à la notion du PPI .....	29
3.3	Rappel de quelques notions utilisées dans la modélisation.....	33
3.3.1	Les contrats futurs .....	33
3.3.2	Modèle de diffusion et processus stochastique .....	34
3.3.3	Décomposition de Cholesky : les nombres aléatoires multi variés. ....	38
3.4	Le processus de modélisation.....	41
3.5	Le modèle de prévision .....	45
3.5.1	Structure du modèle de prévision.....	45
CHAPITRE 4 ANALYSE DES DONNÉES ET PRÉSENTATION DES RÉSULTATS .....		50
4.1	Décomposition de la variance du WPU10. ....	50
4.2	Identification des indicateurs avancés des sous composants du WPU10.....	52
4.2.1	Un exemple d'analyse : le cas du WPU101211 ou Carbon steel scrap.....	53
4.2.2	Tableau récapitulatif des analyses par sous composants du WPU10.....	55
4.3	Test de la robustesse de l'indice synthétique : test historique.....	59
4.4	Un exemple détaillé de prévision : Janvier 2010 à Mars 2010 .....	60
4.4.1	Initialisation : pourvoir notre modèle de données et instructions initiales.....	60
4.4.2	Génération des nombres aléatoires.....	62
4.4.3	Statistiques des nombres aléatoires .....	63

4.4.4	Calcul des taux de croissance mensuels selon les différents scénarios.....	64
4.4.5	Prévision des différentes commodités.....	65
4.4.6	Prévision du WPU10.....	66
4.5	Les résultats de la prévision longue Janvier 2010 à Juin 2010 .....	68
4.5.1	Prévision du WPU10 selon 5 scénarios.....	68
4.6	Conclusion.....	70
CHAPITRE 5 MODÈLE D'ANALYSE DE SÉRIES TEMPORELLES .....		71
5.1	Rappel sur la théorie.....	71
5.1.1	Le modèle.....	71
5.1.2	Importance de la stationnarité dans ce type de modélisation .....	72
5.2	Calibrage du modèle auto régressif au cas du WPU10 .....	73
5.2.1	Analyse initiale.....	73
5.2.2	Estimation du modèle auto régressif .....	78
5.3	Prévision du modèle Auto régressif .....	79
5.3.1	Calibrage du modèle auto régressif d'ordre 1 .....	79
5.3.2	Test historique du modèle .....	80
5.3.3	Prévision sans changement de régime avec une partie hors échantillon.....	81
CONCLUSION .....		82
BIBLIOGRAPHIE .....		85
ANNEXES .....		87

## LISTE DES TABLEAUX

Tableau 2-1 Composition de l'indice S&P 500 .....	12
Tableau 2-2 Valeur historique du prix hebdomadaire moyen.....	13
Tableau 2-3 Exemple d'application du SRC .....	14
Tableau 2-4 Indices de sensibilité avec la méthode de Sobol .....	21
Tableau 2-5 Écart types d'estimation de la méthode de Sobol .....	21
Tableau 2-6 Variables de l'expérience .....	24
Tableau 2-7 Matrice du Plan d'expérience .....	25
Tableau 2-8 Analyse de la variance .....	26
Tableau 2-9 Sensibilité totale appliquée à la planification d'expérience .....	27
Tableau 3-1 Décomposition des PPI .....	30
Tableau 3-2 Décomposition du WPU10 .....	31
Tableau 3-3 Données du WPU03THRU15 selon le BLS .....	32
Tableau 3-4 Données du WPU10 selon le BLS .....	32
Tableau 3-5 Prix futurs d'un tonne de cuivre .....	34
Tableau 3-6 Prix futurs du cuivre.....	36
Tableau 3-7 Taux de croissance mensuel du prix du cuivre basé sur le prix futur .....	36
Tableau 3-8 Exemple d'application du processus de diffusion .....	37
Tableau 4-1 Sous-composants du WPU10 les plus influents dans la contribution à sa variance. .	51
Tableau 4-2 Sous-composants du WPU10 les moins influents sur sa variance.....	51
Tableau 4-3 Analyse du WPU101211.....	53
Tableau 4-4 Analyse des sous-composants de WPU10 .....	56
Tableau 4-5 Les sous-composants de WPU10 regroupés par indicateur .....	57
Tableau 4-6 Matrice de corrélation des prix des matières premières.....	60

Tableau 4-7 Matrice de corrélation après application de la décomposition de Cholesky .....	61
Tableau 4-8 Prix commodités .....	61
Tableau 4-9 Taux de croissance ( $\mu$ ) des prix des commodités.....	61
Tableau 4-10 Variance commodités.....	61
Tableau 4-11 Statistiques nombres aléatoires multi variés .....	64
Tableau 4-12 Taux de croissance mensuels par commodité par scénario.....	65
Tableau 4-13 Prévision des commodités.....	66
Tableau 4-14 Multiplicateur de croissance du WPU10 .....	67
Tableau 4-15 Prévision de l'indice WPU10 .....	67
Tableau 4-16 Prévision du WPU10.....	69
Tableau 5-1 Régression de la première différence du log de WPU10 .....	79
Tableau 5-2 Sous-composants influents du WPU10 et les matières premières qui les impactent .	82
Tableau 5-3 Tableau comparatif des deux modèles de prévision .....	84

## LISTE DES FIGURES

Figure 2-1 Représentation du modèle selon les 3 variables .....	20
Figure 2-2 Exemple de processus pour la planification d'expérience .....	22
Figure 2-3 Synthèse des méthodes d'analyse de sensibilité.....	28
Figure 2-4 Diagramme de décision pour le choix de la méthode d'analyse de sensibilité .....	28
Figure 3-1 Prix futur d'une tonne de cuivre sur douze mois.....	34
Figure 3-2 Intervalle de confiance du prix futur d'une tonne de cuivre sur 12 mois .....	38
Figure 3-3 Diagramme du processus de modélisation .....	42
Figure 3-4 Schématisation du processus de prévision .....	45
Figure 3-5 Schématisation des résultats de la prévision sur un intervalle de temps $[t, t+n]$ .....	46
Figure 3-6 Schématisation des résultats de la simulation pour un actif .....	48
Figure 3-7 Recomposition du WPU10 selon le poids de chaque métal dans l'indice .....	49
Figure 4-1 Distribution normale des résidus de la régression du WPU101211 sur le HRC .....	54
Figure 4-2 Répartition des résidus de la régression du WPU101211 sur le HRC .....	54
Figure 4-3 Comparaison des valeurs réelles et prédites du WPU101211 .....	55
Figure 4-4 Test historique de la robustesse de l'indice synthétique .....	59
Figure 4-5 Prévision de l'indice WPU10.....	68
Figure 4-6 Prévision du WPU10.....	69
Figure 5-1 Données historiques du WPU10.....	73
Figure 5-2 Oscillation du WPU10 par rapport à sa tendance.....	74
Figure 5-3 Auto corrélations des résidus de la régression de WPU10 par rapport à sa tendance..	75
Figure 5-4 Résidus de la régression de WPU10 Tendance et saisonnalité .....	77
Figure 5-5 Auto corrélations des résidus de la régression de WPU10 Tendance et saisonnalité ..	77
Figure 5-6 Test historique du WPU10 selon un modèle auto régressif .....	80
Figure 5-7 Prévision du WPU10 selon un modèle auto régressif .....	81

**LISTE DES SIGLES ET ABRÉVIATIONS**

$X, Y$	Variables aléatoires
$X_i$	i-ème composante de $X$
$E(X)$	Espérance de $X$
$V(X)$	Variance de $X$
$\sigma_X$	Écart type de $X$
$COV(X, Y)$	Covariance de $X$ et $Y$
$\rho_{X,Y}$	Coefficient de corrélation entre $X$ et $Y$
$N(\mu, \Sigma)$	Loi normale de moyenne $\mu$ et de variance $\Sigma$

## **LISTE DES ANNEXES**

ANNEXE 1 – Feuille Excel d’initialisation de la simulation

ANNEXE 2 – Statistiques des nombres aléatoires multivariés

ANNEXE 3 – Volatilité et taux de croissance mensuel par scénario

ANNEXE 4 – Prévision

ANNEXE 5 – Détail des prévisions

ANNEXE 6 – Code VBA

## INTRODUCTION

Les scientifiques sont souvent confrontés à une difficulté qui est celle de pouvoir reconstruire des phénomènes observés à l'aide d'outils mathématiques. Dans cette quête de modélisation afin de répliquer les phénomènes naturels, une des manières de faire est d'expliquer les phénomènes et notamment leurs comportements en faisant des liens et des analogies avec les comportements d'autres phénomènes que l'on connaît mieux : c'est cette pratique que l'on nomme analyse de sensibilité. D'une manière générique, selon une dualité mathématique intrants/extrant d'un processus, l'analyse de sensibilité est l'étude des relations entre les variations de l'extrant d'un modèle mathématique et les variations provenant des intrants. Ces variations peuvent être le fait de variations naturelles intrinsèques au processus comme des processus stochastiques, des erreurs de mesure, des absences d'information, une mauvaise ou faible compréhension des mécanismes primordiaux sous-jacents au processus que l'on cherche à expliquer.

Aussi dans le cadre du processus de modélisation, l'analyse de sensibilité peut permettre entre autres :

- i. de juger de la fidélité du modèle par rapport au phénomène observé dans la réalité. Ainsi, si l'analyse exhibe une importance forte d'une variable d'entrée qui en réalité est connue comme non influente, le modèle ne reflétera pas correctement le processus. Il sera alors nécessaire de modifier le modèle.
- ii. d'identifier les variables influentes sur la variabilité de la réponse du modèle et, le cas échéant, diminuer cette volatilité en améliorant la qualité des intrants ou en modifiant la structure du modèle pour atténuer les effets identifiés. L'identification des variables moins influentes nous permettra d'alléger le modèle étudié et, le cas échéant de les supprimer du modèle.
- iii. d'identifier les variables influentes et leurs interactions et de connaître le phénomène étudié et ses mécanismes sous-jacents.

Ainsi, dans le domaine financier par exemple, une bonne gestion passe par la capacité à pouvoir budgétiser et réadapter les stratégies de génération de revenus face aux changements. Cela est d'autant plus vrai que les paramètres impactant les marchés sont aussi divers que les acteurs qui y interviennent. Dans ce contexte, comprendre et budgétiser, c'est-à-dire prévoir les revenus et les



dépenses, revient à les modéliser et poser des hypothèses sur l'évolution des paramètres pouvant les impacter et mieux les contrôler.

Dans ce cadre, les entreprises manufacturières sont confrontées à la difficulté d'établir des prévisions robustes de leurs budgets. Ainsi, la volatilité des prix des matières premières qui constituent une grande partie de leurs dépenses est de plus en plus difficile à établir. Pour ce faire, dans le cadre de ce mémoire on s'est intéressé à un indice représentant le prix des métaux publié par le bureau américain de la statistique du travail comme indicateur de l'évolution du prix de ces commodités.

Ainsi, après avoir fait un état de l'art dans le domaine de l'analyse de sensibilité dans le chapitre 1 et présenté les différentes méthodes et indices de sensibilité dans le chapitre 2, on présentera une application de l'analyse de sensibilité dans un cadre de prévision dans le chapitre 3. Dans le chapitre 4, on verra comment la connaissance de la propagation de la variabilité dans les prix des métaux nous permet d'estimer un futur probable. Enfin, dans le chapitre 5, le modèle de diffusion est comparé au modèle autorégressif utilisé dans l'analyse de série temporelle dont fait partie le WPU10 dans un cadre de prévision pour pouvoir tester son efficacité dans un contexte qui est celui du marché financier.

## CHAPITRE 1 REVUE DE LITTÉRATURE

### 1.1 Introduction

Dans ce chapitre nous présentons une revue de littérature de l'analyse de sensibilité, nous passons en revue notamment les différents types d'analyses de sensibilité.

Soit  $P$  un processus décrivant un phénomène que l'on veut étudier à l'aide d'un modèle mathématique. Ce processus contient deux catégories de variables selon leurs rôles. Un premier groupe de variables  $X$  que l'on nommera intrants du processus et un deuxième groupe de variables mesurées ou observées à la sortie du processus et considérées comme réponse ou extrant du processus représenté par la variable  $Y$ . Le diagramme ci-dessous donne une représentation schématique d'un tel processus :



où  $X$  est un ensemble de variables intrants au modèle et  $Y$  une variable extrant ou réponse du modèle.

Effectuer une analyse de sensibilité de  $Y$  par rapport aux  $X$  revient à étudier dans quelle mesure les perturbations sur  $X$  affectent les valeurs de  $Y$ . On cherche donc à établir un lien entre les variabilités de  $X$  et  $Y$ . Cette analyse de sensibilité se fait selon trois types de techniques : les méthodes de criblage, l'analyse de sensibilité locale et l'analyse de sensibilité globale.

Les méthodes de criblage sont entre autres présentées par Campolongo [2], elles abordent l'analyse de sensibilité en classifiant les variables d'entrée par ordre d'impact sur la variable réponse du modèle. Les analyses de sensibilité locale et globale, présentées entre autres par Saltelli et al. [11, 15, 17], sont des méthodes d'analyses quantitatives, qui en plus d'établir une hiérarchie au sein des variables d'entrée, donnent un ordre de grandeur des écarts au sein de cette hiérarchie.

Dans les prochains paragraphes nous présenterons chacune de ces méthodes ainsi que leurs avantages et inconvénients.

## 1.2 Méthodes de criblage

Les méthodes de criblage servent d'outils d'exploration du comportement des sorties d'un processus en présence d'un très grand nombre de variables d'entrée. Elles font appel à la discrétisation des entrées en plusieurs valeurs nommées niveaux ou modalités, ce qui en fait des méthodes dites déterministes car elles n'utilisent pas de lois de probabilité sur les entrées. La particularité de ces modèles est la facilité et la rapidité de leurs applications, car elles requièrent moins de temps en termes de modélisation ou de calcul.

### 1.2.1 Méthode de criblage simple ou à très grande dimension

Les techniques qui font partie de cette catégorie nous sont pour la plupart fournies par la pratique des plans d'expériences et sont parmi les plus facilement applicables dans la mesure où elles posent peu d'hypothèses contraignantes quand à la nature du modèle. Parmi ces hypothèses on trouve notamment l'indépendance en probabilité des variables d'entrées du modèle, la monotonie de la variation de la sortie du modèle par rapport à chaque variable d'entrée et la faiblesse du nombre de variables d'entrées influentes par rapport au nombre total de variables d'entrées. Bertrand Iooss [5] nous présente ces méthodes :

- i. **Criblage avec plan supersaturé** : à cet effet effectuons d'abord un bref rappel sur une propriété des plans d'expériences. En effet dans la pratique des plans d'expériences, il est souvent commun d'avoir un nombre d'observations plus grand que le nombre de variables à étudier, ce qui nous permet notamment d'avoir des estimés en termes de probabilités des coefficients représentant l'influence (effet) des variables entrantes sur les variables de sorties; ces effets pouvant être regroupés en plusieurs catégories : effet principal, effet d'interaction (ou second ordre), etc. qui peuvent être analysés avec un **plan factoriel complet**. Cependant, dans beaucoup de situations spécifiques, avoir des observations suffisantes est difficile voir même impossible à cause d'un manque de ressources. Dans ce contexte les plans supersaturés ont été introduits pour pallier à cette difficulté. Dans le cas des plans supersaturés nous avons affaire à la même situation que celle de la résolution d'un système d'équations à plusieurs inconnues sans avoir un nombre suffisant d'équations. Pour résoudre un tel système, on fera donc appel

aux diverses méthodes de résolution de système d'équations, que l'on n'abordera pas dans le cadre de ce mémoire.

- ii. **Criblage par groupe :** Il est plus adapté aux expérimentations numériques et consiste en la division des variables d'entrées en plusieurs groupes et à identifier les plus influentes. Le but étant de multiplier les compositions des groupes en éliminant les groupes moins influents jusqu'à l'obtention des groupes tous influents. Les étapes de la démarche sont:

- 1) Les variables indépendantes sont regroupées en groupes pas nécessairement égaux, sur la base de l'intuition du chercheur. Chaque groupe est alors étudié comme un facteur à deux niveaux :
  - le facteur-groupe est au niveau (+) lorsque tous les facteurs du groupe sont au niveau (+).
  - le facteur-groupe est au niveau (-) lorsque tous les facteurs du groupe sont au niveau (-).
- 2) Après l'identification des groupes influents, les non influents sont éliminés puis dans une seconde phase, soit chaque variable dans les groupes influents est testée individuellement ou alors les groupes sont divisés en de nouveaux groupes plus petits pour être testés à nouveau jusqu'à l'obtention des variables les plus influentes.

Cette technique nécessite cependant une connaissance préalable du sens de variation de la sortie en fonction du sens de variation de chaque entrée, connaissance qui n'est pas toujours disponible.

- iii. **Criblage par bifurcations séquentielles :** C'est une méthode de criblage par groupe avec deux groupes. Contrairement aux autres méthodes par groupes elle n'utilise pas de plan d'expérience de tamisage et des méthodes d'analyse de régression ou d'analyse de variance. C'est une approche similaire au principe de la recherche dichotomique : les expériences sont choisies séquentiellement en utilisant les résultats des expériences précédentes pour orienter la recherche des facteurs influents. Comme pour le criblage par groupe, son coût est fonction du

nombre de variables influentes et de la stratégie de classement et la capacité à identifier les entrées influentes afin de les rassembler au sein d'un même groupe.

### 1.2.2 La méthode de Morris

Les méthodes de criblage présentées par Campolongo [2] ont été développées par Morris en 1991. Elles reposent sur un principe simple à savoir qu'un modèle comportant un nombre élevé de facteurs est difficile à explorer mais souvent quelques facteurs seulement sont influents. Par conséquent, on cherchera dans le cadre d'un objectif qualitatif à identifier rapidement ces facteurs. La méthode de Morris, permet de classer les facteurs en trois groupes selon leurs effets :

- effets négligeables
- effets linéaires et sans interaction
- effets non linéaires avec interactions

Pour ce faire, Morris introduit un premier ajustement à la méthode commune des dérivées en utilisant un pas de perturbations des variables très grand par opposition aux dérivées dont le pas est infiniment petit. Il introduit dans ce sens le concept d'effet élémentaire qui est le changement dans la valeur de l'extrait généré par la perturbation. Par ailleurs il introduit un second ajustement par rapport aux dérivées en calculant plusieurs fois cet effet élémentaire pour la variable d'entrée considérée puis on calcule une moyenne de ces effets élémentaires. Par opposition aux méthodes communes de dérivées, celle de Morris fait le calcul de la dérivée à plusieurs endroits et en prend une moyenne plutôt que de se concentrer sur un seul point pour tirer une conclusion.

## 1.3 Analyse de sensibilité locale

La technique d'analyse de sensibilité locale étudie essentiellement les variations de notre variable de sortie  $Y$  pour des petites variations de nos variables d'entrée  $X$  autour de valeurs locales  $x_i^0$ . Mathématiquement, l'analyse de sensibilité locale correspond au calcul des dérivées de la variable  $Y$  par rapport aux variables  $X$ . L'approche basée sur les dérivées à l'avantage d'être très efficace en temps de calcul sur ordinateur. Cependant, elle requiert beaucoup plus d'implications de la part de l'analyste qui doit préparer et coder l'algorithme. Comme le fait remarquer Saltelli [15], le problème fatal avec cette méthode c'est qu'elle est non efficace

lorsque les intrants sont incertains ou que la linéarité du modèle est inconnue. Ainsi, si les dérivées sont efficaces localement, elles le sont moins dans le cas de l'exploration d'un espace de solution surtout si le modèle est non linéaire. Pour le cas des modèles linéaires on peut extrapoler la propriété d'un point éloigné en fonction de la dérivée localement étudiée. A cet égard, Julien Jacques [6] fait remarquer que la technique d'analyse de sensibilité locale est souvent un facteur à la fois et de ce fait, elle se limite aux alentours de la valeur d'un point d'intérêt particulier noté  $x_i^0$  calculant les différentes perturbations représentées par l'indice :

$$S_i = \frac{\delta y}{\delta x_i}(x_1^0, \dots, x_k^0) \quad (1.2)$$

Ainsi Turyani et al. [18] regroupent les avantages et inconvénients de l'analyse de sensibilité locale en 3 groupes :

- **analyse et calibrage de modèle** : on peut avoir recours à une analyse locale pour étudier les effets des variables explicatives sur la variable réponse d'une part, et d'autre part, pour calibrer le modèle en identifiant un nombre maximum de variables pertinentes.
- **analyse d'incertitude** : on reconnaît que, dans le cas de l'étude de la propagation des erreurs, l'analyse de sensibilité locale est moins efficace comparativement à l'analyse de sensibilité globale; cependant on pense que son application est beaucoup plus aisée dans bien des cas.
- **estimation des paramètres** : on retrouve souvent dans les méthodes d'estimation de paramètres lors de régressions des estimations de coefficients de sensibilité ce qui facilite d'avance l'analyse de sensibilité locale.

## 1.4 Analyse de sensibilité globale

Afin de donner une définition claire de l'analyse de sensibilité globale, on peut se référer à celle donnée par Julien Jacques [6] : « **l'analyse de sensibilité globale** s'intéresse quant à elle à la variabilité de la sortie du modèle dans son domaine de variation. Elle étudie comment la variabilité des entrées se répercute sur celle de la sortie, en déterminant quelle part de variance de la sortie est due à telles entrées ou tel ensemble d'entrées. Il est possible de distinguer l'analyse locale de l'analyse globale de la sorte : l'analyse locale s'intéresse à la valeur de la réponse,

tandis que l'analyse globale s'intéresse à sa variabilité. ». Tout comme dans le cas de l'analyse de sensibilité locale où on utilise la dérivée pour déterminer une relation entre les variables entrantes et sortantes du modèle, on peut également, dans le cadre de l'analyse de sensibilité globale utiliser des indices de sensibilité selon la nature du modèle. Dans le cadre de ce mémoire nous nous intéressons particulièrement aux différentes méthodes (Sobol par exemple) et indices relevant de ce type d'analyse que nous développerons dans le prochain chapitre.

## CHAPITRE 2      MÉTHODES ET INDICES DE L'ANALYSE DE SENSIBILITÉ

Ce chapitre vise à présenter une revue des différentes méthodes et indices utilisés en analyse de sensibilité. Dans le contexte d'une analyse de sensibilité on dispose généralement d'une variable extrant  $Y$  et de plusieurs variables intrants  $X_1, X_2, \dots, X_p$ . L'évaluation de la sensibilité de  $Y$  par rapport à chacune des variables intrants  $X_i$ , ( $i = 1, 2, \dots, p$ ), se fait à l'aide d'indices de sensibilité. La méthode d'évaluation de ces indices dépend de la nature du modèle liant  $Y$  à  $X_i$ .

### 2.1 Indices de sensibilité

#### 2.1.1 Indices de sensibilité dans le cas du modèle linéaire

Les coefficients de corrélation sont des cas classiques d'indices de sensibilité conditionné par la linéarité ou la monotonie du modèle, cependant il existe également d'autres types d'indices de sensibilité.

Soit le modèle linéaire 
$$Y = \beta_0 + \sum_{i=1}^p \beta_i X_i \quad (2.1)$$

$Y$  représente ici un indice composite des variables  $X_i$  selon les coefficients  $\beta_i$  qui dans bien des cas peuvent être de coefficients de pondérations (indices boursiers, portefeuille financier, etc.) mais aussi peuvent être des coefficients de régression. Par ailleurs, afin de faciliter nos démonstrations, on supposera les  $X_i$  comme des variables aléatoires indépendantes. Ce qui nous permet d'écrire l'espérance  $E(Y)$  et la variance  $V(Y)$  selon les équations (2.2) et (2.3) :

$$E(Y) = \beta_0 + \sum_{i=1}^p \beta_i E(X_i) \quad (2.2)$$

$$V(Y) = \sum_{i=1}^p \beta_i^2 V(X_i) \quad (2.3)$$

où  $E(X_i)$  et  $V(X_i)$  désignent respectivement la moyenne et la variance des  $X_i$ ,  $i = 1, \dots, p$ .



Cette écriture nous permettant d'identifier la contribution de la variance de chaque  $X_i$  dans la variance totale de Y par  $\beta_i^2 V(X_i)$ .

**Définition 2.1 :** l'indice de sensibilité global nommé *SRC (Standardized Regression Coefficient)* représentant la part de variance de la réponse Y due à la variance de la variable  $X_i$  est défini par

$$SRC_i = \frac{\beta_i^2 V(X_i)}{V(Y)} = \beta_i^2 \frac{\sigma_{X_i}^2}{\sigma_Y^2} \quad (2.4)$$

Où  $\sigma_{X_i}^2$  et  $\sigma_Y^2$  représentent la variance de  $X_i$  et celle de Y.

L'indice ainsi défini apporte, contrairement au coefficient de corrélation, l'information quand à l'impact de chaque variable intrant  $X_i$  sur la réponse Y sans se préoccuper du signe de l'impact.

Par ailleurs, on note :

$$SRC_i = (S_{X_i}^\sigma)^2 = \left(\frac{\sigma_{X_i}}{\sigma_Y} \beta_i\right)^2 \quad (2.5)$$

où l'indice  $S_{X_i}^\sigma$  ainsi introduit est équivalent au coefficient de corrélation linéaire de Pearson  $\rho_{X_i,Y}$  entre la réponse du modèle et les variables d'entrées  $X_i$ .

En effet, on sait que :

$$\rho_{X_i,Y} = \frac{Cov(Y, X_i)}{\sigma_{X_i} \sigma_Y}$$

Alors

$$\rho_{X_i,Y} = \frac{\sigma_{X_i}}{\sigma_Y} \beta_i = S_{X_i}^\sigma \quad \text{car } Cov(X_i, Y) = \beta_i \sigma_{X_i}^2 \quad (2.6)$$

*Montrons que*  $Cov(X_i, Y) = \beta_i \sigma_{X_i}^2$

On sait que les variables  $X_i$  sont indépendantes par conséquent étudier la covariance entre n'importe quel  $X_i$  et Y revient à étudier la covariance entre  $X_i$  et  $\beta_i X_i$ .

En effet on sait que :

$$COV(U+V, W) = COV(U, W) + COV(V, W)$$

Pour toutes variables U, V et W.

Donc  $Cov(X_i, Y)$  s'écrit :

$$Cov(X_i, Y) = Cov(X_i, \beta_1 X_1) + \dots + Cov(X_i, \beta_{i-1} X_{i-1}) + Cov(X_i, \beta_i X_i) + \dots + Cov(X_i, \beta_p X_p)$$

Or comme les  $X_i$  sont indépendants alors toutes les covariances de l'équation ci-dessus sont nulles à l'exception de  $Cov(X_i, \beta_i X_i)$ .

Ce qui nous permet d'écrire que :

$$Cov(X_i, Y) = Cov(X_i, \beta_i X_i)$$

Par ailleurs, on sait que :

- $Cov(aH, Z) = aCov(H, Z)$
- $Cov(Z, Z) = Var(Z)$ , pour toutes variables H et Z et toute constante a.

Par conséquent,  $Cov(X_i, Y) = \beta_i \sigma_{X_i}^2$

Afin de régler le problème d'intercorrélations entre les variables explicatives  $X_i$  qui souvent expliquent l'existence d'une corrélation entre une variable explicative et la variable réponse par un effet de transition de corrélation, **l'indice de corrélation partiel de  $X_i$**  a été introduit, il est défini par

$$PCC(X_i) = \rho_{Y, X_i | X \sim i} = \frac{COV(Y, X_i | X \sim i)}{\sqrt{V(Y | X \sim i) V(X_i | X \sim i)}} \quad (2.7)$$

où  $X \sim i$  est le vecteur X des variables explicatives privé de sa i-ème composante. Tout comme le  $SRC_i$ , le  $PCC_i$  sert également à déterminer la part de la variable  $X_i$  dans la variance totale de Y. Cependant, contrairement au  $SRC_i$ , il peut être négatif c'est pourquoi on prendra la valeur absolue de celui-ci. Les  $PCC_i$  permettent d'éliminer l'influence des autres variables et sont donc adaptés au cas où les variables d'entrée sont corrélées (Saltelli et al. [14]).

### *Exemple*

Nous avons choisi de présenter une application des indices ci-dessus dans le cadre de l'analyse d'un indice phare de l'économie et spécifiquement de l'économie américaine à savoir le S&P 500. Il est de notoriété que cet indice permet de prendre le pouls de l'économie américaine dans la mesure où il est constitué essentiellement des grandes entreprises de cette économie.

Les 10 compagnies les plus importantes composants l'indice S&P 500					
Pays	Compagnie	Valeur boursière (Million \$)	Poids dans l'indice S&P 500	Poids dans le secteur	Secteur
USA	Exxon Mobil (ESSO)	368 711 ,8	<b>3,23%</b>	26,81%	Energie
USA	Apple Inc.	295 888,5	<b>2,59%</b>	13,88%	Technologies de l'information
USA	Microsoft	210 205,8	<b>1,84%</b>	9,86%	Technologies de l'information
USA	General Electric Co. (GE)	194 874,8	<b>1,70%</b>	15,57%	Industries
USA	Chevron Corp.	183 634,1	<b>1,61%</b>	13,35%	Energie
USA	IBM	182 328,9	<b>1,60%</b>	8,56%	Technologies de l'information
USA	Procter & Gamble	180 072,1	<b>1,58%</b>	14 ,82%	Consommation
USA	AT&T Inc.	173 635,8	<b>1,52%</b>	48,91%	Télécommunications
USA	Johnson & Johnson	169 855,8	<b>1,49%</b>	13,62%	Santé
USA	JP Morgan Chase & Co	165 827,5	<b>1,45%</b>	9,03%	Finance

Tableau 2-1 Composition de l'indice S&P 500

Dans le tableau ci-dessus tiré du « S&P 500, Index Fact Sheet » on peut voir les 10 plus grandes entreprises qui composent l'indice (et donc l'économie américaine indirectement) en 2010. On peut constater que pour certains comme AT&T et Exxon Mobil ce sont des joueurs majeurs de

leurs industries (poids respectifs dans leurs industries 48,91% et 26,81% vs poids respectifs dans l'indice de 1,52% et 3,23%). Le poids dans le S&P 500 dépend surtout de la valeur boursière de chaque entreprise. Aussi le poids d'Exxon (3,23% de l'économie américaine) dans le S&P 500 dépasse de loin celui d'AT&T parce que la valeur boursière d'Exxon (deuxième plus grande de la planète, 368.711,8 millions) dépasse celle d'AT&T (173.635,8 millions).

Afin de procéder à une étude de la variabilité du S&P500 et notamment la compréhension de l'impact sur l'économie de la variabilité du prix de chaque entreprise nous avons étudié l'impact de la variation du prix unitaire de chaque action de trois entreprises sur la variation du S&P500 sur une période allant du 1<sup>er</sup> janvier 2007 au 31 décembre 2010.

Tableau 2-2 Valeur historique du prix hebdomadaire moyen

Moyenne hebdomadaire de la valeur boursière				
Date	S&P 500 (Valeur en points)	General Electric (Valeur en \$ US)	EXXON (Valeur en \$ US)	Apple (Valeur en \$ US)
27-déc.-10	1254,66	17,92	72,81	1254,66
20-déc.-10	1245,76	17,77	72,24	1245,76
13-déc.-10	1242,52	17,89	72,41	1242,52
6-déc.-10	1223,87	16,70	71,20	1223,87
29-nov.-10	1189,08	15,73	68,86	1189,08
22-nov.-10	1198,07	16,14	70,10	1198,07
15-nov.-10	1200,44	16,23	71,01	1200,44
8-nov.-10	1223,24	16,62	69,49	1223,24
1-nov.-10	1185,71	16,09	66,72	1185,71
...	...	...	...	...
5-févr.-07	1448,33	36,27	75,65	1448,33
29-janv.-07	1422,03	36,07	73,65	1422,03
22-janv.-07	1430,47	37,13	73,65	1430,47
16-janv.-07	1430,73	38,00	72,50	1430,73
8-janv.-07	1409,26	37,45	73,88	1409,26
3-janv.-07	1418,03	37,41	76,26	1418,03

Après calcul du taux de croissance hebdomadaire du prix de chaque action de chaque entreprise tel que donné ci-dessus, on peut calculer l'écart type de ce taux de croissance et en appliquant l'équation (2.4), on peut identifier la part de la variance du S&P 500 liée au prix de l'action de chaque entreprise :

Tableau 2-3 Exemple d'application du SRC

:

écart type de la croissance continue hebdomadaire			
S&P 500	General Electric	EXXON	Apple
3,3007%	6,0575%	3,4387%	3,3007%
Part de la variance du S&P 500 due à chaque entreprise sur une base hebdomadaire			
GE		EXXON	Apple
0,0973%		0,1132%	0,0671%

On peut donc conclure qu'en moyenne ces 3 entreprises, qui représentent 7,53% de l'économie américaine, ont contribué pour une part de 0,27% à la variance hebdomadaire du S&P500.

### 2.1.2 Indices de sensibilité cas du modèle non linéaire et monotone

Les indices SRC et PCC sont conditionnés par la linéarité du modèle. Cependant, lorsque cette condition n'est plus respectée, mais que le modèle est monotone c'est-à-dire que la variable réponse est monotone par rapport à chacune des variables explicatives lorsque les autres sont fixées, un estimateur non paramétrique du coefficient de corrélation peut être utilisé. En effet, comme on a établi que le SRC est le carré du coefficient de corrélation, dans le cas d'un modèle non linéaire mais monotone on pourra utiliser le coefficient de corrélation basé sur les rangs [5].

Dans ce cas ci, on peut utiliser la transformation de rang pour avoir un coefficient de corrélation estimé de façon non paramétrique en utilisant des techniques de calcul de coefficient de corrélation de **Spearman** ou de **Kendall** par exemple. Par la suite, on calcule un nouvel indice **SRRC** ou « Standardized rank regression coefficient » pour remplacer l'indice SRC. De même, on introduit un indice **PRCC** ou « Partial Rank Correlation Coefficient » pour remplacer l'indice PCC.

Soit une simulation de taille  $n$  du modèle étudié  $Y = f(X_1, X_2, \dots, X_p)$ ,  $p$  entier, telle qu'elle se présente sous la forme :

$$\begin{pmatrix} Y_1 \\ \vdots \\ Y_n \end{pmatrix} = \begin{pmatrix} X_{1,1} & X_{1,2} & X_{1,3} & \dots & X_{1,p} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ X_{n,1} & \vdots & \vdots & \vdots & X_{n,p} \end{pmatrix}$$

où les différents Y sont les extrants et les X de la même ligne sont les intrants ayant servi à les calculer.

À chaque simulation  $i$  de  $1$  à  $n$ , on associe son rang selon une variable. Le rang 1 sera affecté à la simulation qui a la plus petite valeur, et le rang  $n$  à celle qui a la plus grande valeur. Pour trouver le lien entre deux variables  $(Y, X_j)$ , ( $j = 1, \dots, p$ ), on cherchera à comparer les classements selon les rangs issus de ces deux variables.

Dans notre cas, nous créons le vecteur  $R_Y$  des rangs selon la variable Y et les vecteurs  $R_{X_j} = (R_{X_{1,j}}, R_{X_{2,j}}, R_{X_{3,j}}, \dots, R_{X_{n,j}})$  des rangs selon  $X_j$  pour chaque variable d'entrée avec  $j = (1, \dots, p)$ . On pourra ainsi avoir une nouvelle matrice des rangs issue de notre précédente matrice en remplaçant les valeurs de  $(Y_i, X_{i,j})$  avec  $i = (1, \dots, n)$  et  $j = (1, \dots, p)$  par leurs rangs dans chaque colonne de la matrice.

Ce qui nous permettra de calculer un nouveau type de corrélation selon les rangs :

$$\rho_j^S = \rho^S(Y, X_j) = \rho(R_Y, R_{X_j}) \quad (2.8)$$

❖ le coefficient de **régression standard** sur les rangs noté SRRC ( $\cdot, \cdot$ ) équivalent du SRC de l'équation (2.4) basé sur le couple  $(R_Y, R_{X_j})$ , s'écrit :

$$SRRC_j = SRRC(Y, X_j) = SRC(R_Y, R_{X_j}) \quad (2.9)$$

❖ le coefficient de **corrélation partiel** sur les rangs noté PRCC ( $\cdot, \cdot$ ), équivalent du PCC de l'équation (2.7) sur le couple  $(R_Y, R_{X_j})$ , s'écrit :

$$PRCC_j = PRCC(Y, X_j) = PCC(R_Y, R_{X_j}) \quad (2.10)$$

### 2.1.3 Indices de sensibilité cas du modèle non linéaire et non monotone

Soit un modèle générique du type :

$$Y = f(X_1, X_2, \dots, X_p)$$

Les  $X_i$  ont une variabilité limitée et on souhaite étudier l'impact de la variable  $X_i$  sur la variance de Y en fixant un  $X_i$  à une valeur  $X_i^*$  telle que

$$V(Y|X_i = X_i^*)$$

est la variance conditionnelle de  $Y$  lorsque  $X_i$  est fixé à une valeur  $X_i^*$ . On peut cependant assumer qu'ayant fixé une variable, la variance conditionnelle  $V(Y|X_i = X_i^*)$  serait inférieure à la variance totale  $V(Y)$  de  $Y$ . Par ailleurs  $V(Y|X_i = X_i^*)$  devient une mesure de l'importance relative de  $X_i$ . Ainsi, plus la variance conditionnelle est petite plus la variable  $X_i$  est importante et vice versa. Cependant cette approche est peu pratique car cela rend l'analyse de sensibilité dépendante du point  $X_i^*$  où l'on fixe  $X_i$ . Une manière d'y remédier est d'utiliser plutôt l'espérance  $E[V(Y|X_i)]$ . Nous introduisons pour cela le théorème suivant pour en montrer le lien.

**Théorème 2.1 : Théorème de la variance totale**

Soit un couple  $(X, Y)$  de variables aléatoires, où  $Y$  prend ses valeurs dans  $R$  et  $X$  dans un ensemble fini ou dénombrable, ou dans  $R$ . Si la variance de  $Y$  est finie, alors :

$$V(Y) = V(E[Y|X]) + E[V(Y|X)] \quad (2.11)$$

***Démonstration du théorème de la variance totale***

Par définition :

$$V(Y) = E(Y^2) - [E(Y)]^2$$

et on sait que

$$E[E(Y|X)] = E[Y]$$

On peut donc observer deux choses :

d'une part,

$$V(Y|X) = E(Y^2|X) - [E(Y|X)]^2$$

soit donc,

$$\begin{aligned} E[V(Y|X)] &= E[E(Y^2|X)] - E[[E(Y|X)]^2] \\ &= E(Y^2) - E[[E(Y|X)]^2] \end{aligned}$$

d'autre part, parce que

$$E[E(Y|X)] = E[Y]$$

on peut écrire

$$V[E(Y|X)] = E[[E(Y|X)]^2] - [E(Y)]^2$$

Aussi, en sommant les écritures de  $E[V(Y|X)]$  et  $V[E(Y|X)]$  observées aux deux étapes on obtient la variance totale de  $Y$ , tel que :

$$V(Y) = V(E[Y|X]) + E[V(Y|X)]$$

Ainsi, si l'on considère l'espérance  $E(V(Y|X_i))$  pour tous les  $X_i^*$  possibles pour chaque intrant cela élimine le problème de la dépendance par rapport au  $X_i^*$ . Cette mesure  $E(V(Y|X_i))$  sera toujours inférieure ou égale à  $V(Y)$ .

On peut donc écrire :  $V(Y) = V(E[Y|X_i]) + E[V(Y|X_i)]$

Dans l'éventualité où on aura fixé  $X_i$  dans les  $X$ , si l'on observe une diminution de l'espérance de la variance de  $Y$  ou une augmentation de la variance de l'espérance de  $Y$ , on pourra dire que le  $X_i$  fixé est influent.

Par ailleurs, selon l'équation (2.11), la variance  $V(E(Y|X_i)) \leq V(Y)$

**Définition 1.2 :** l'indice de sensibilité de premier ordre de  $X_i$  par rapport à  $Y$ , est l'indice basé sur la variance conditionnelle  $V(E(Y|X_i))$  et appelé effet de premier ordre de  $X_i$  sur  $Y$  noté  $S_i$ , tel que :

$$S_i = \frac{V(E(Y|X_i))}{V(Y)} \quad (2.12)$$

$S_i$  est toujours compris entre 0 et 1 et mesure l'importance de la variable  $X_i$  sur  $Y$ .

Lorsque le modèle est linéaire, alors  $S_i$  correspond au SRC (voir équation (2.4)). En effet :

$$V(E[Y|X_i]) = V(E[\beta_0 + \sum_{i=1}^p \beta_i X_i | X_i]) = V(\beta_i X_i) = \beta_i^2 V(X_i) \quad (2.13)$$

La sensibilité de premier ordre a été introduite par Sobol [16] dans un cadre beaucoup plus exhaustif où la variance d'une variable réponse devient une combinaison de variance d'effets principaux et de plusieurs effets interactions. Le sujet a été traité par Antoniadis [1], Kaplan, Kaya et al. [8], Julien Jacques [6], Tarantola [17], et Saltelli [15].

Ainsi si nous reprenons notre modèle précédent tel que l'on a :



$$Y = f(X_1, X_2, \dots, X_p) \text{ , avec } p \in \mathbb{N}. \quad (2.14)$$

Sobol établit un théorème de la décomposition de la variance analogue à celui de l'ANOVA.

**Théorème 2.2 : Décomposition de la variance de Sobol**

Lorsque les variables intrants  $X_i$  sont indépendants, la décomposition de la variance du modèle (2.14) s'établit comme :

$$V(Y) = \sum_{i=1}^p V_i + \sum_i \sum_j V_{ij} + \sum_i \sum_j \sum_p V_{ijp} + \dots + V_{1,2,3,4,\dots,p} \quad (2.15)$$

Avec

$$\begin{aligned} V_i &= V[E(Y|X_i)] \\ V_{ij} &= V[E(Y|X_i, X_j)] - V_i - V_j \\ V_{ijp} &= V[E(Y|X_i, X_j, X_p)] - V_{ij} - V_{ip} - V_{jp} - V_i - V_j - V_p \end{aligned}$$

Cela nous permet de reconnaître un indice de sensibilité précédemment identifié en (2.12) comme indice de sensibilité de premier ordre, mais aussi d'introduire des indices de sensibilité d'ordre supérieur qui reflètent les interactions entre les variables intrants dans la variance totale de la réponse Y.

En effet en divisant (2.15), par  $V(Y)$  on obtient l'expression :

$$1 = \sum_{i=1}^p S_i + \sum_i \sum_j S_{ij} + \sum_i \sum_j \sum_p S_{ijp} + \dots + S_{1,2,3,4,\dots,p} \quad (2.16)$$

Avec nos effets de premier ordre de toutes les variables intrants regroupés sous :

$$S_i = \frac{V_i}{V(Y)} = \frac{V[E(Y|X_i)]}{V(Y)}$$

Ainsi pour  $p = 4$  par exemple, on aura :

$$S_1 + S_2 + S_3 + S_4 + S_{12} + S_{13} + S_{14} + S_{23} + S_{24} + S_{34} + S_{123} + S_{124} + S_{134} + S_{234} + S_{1234} = 1$$

Avec  $S_1, S_2, S_3, S_4$  comme indices de premier ordre.

L'interprétation de ces indices est facile. En effet on sait que tous les indices sont positifs d'une part et d'autre part d'après (2.16) leur somme est égale à 1. Ainsi, plus l'indice sera grand et proche de 1, plus la variable aura d'importance. Le nombre d'indices de sensibilité ainsi

construit, de l'ordre 1 à l'ordre  $p$ , est égal à  $2^p - 1$ . Lorsque le nombre de variables d'entrée  $p$  est grand, le nombre d'indices de sensibilité explose. L'estimation et l'interprétation de tous ces indices deviennent alors difficiles.

Pour pallier en partie à cette difficulté, Saltelli [15] a alors introduit des indices de sensibilité totaux, qui expriment la sensibilité totale de la variance  $Y$  à une variable, en l'occurrence en englobant dans une seule mesure la sensibilité à la variable seule et la sensibilité aux interactions de cette variable avec d'autres variables. Nous présentons ci-dessous ses conclusions telles que montrées par Julien Jacques [6] :

**Définition 1.2 : Indice de sensibilité total**

L'indice de sensibilité total  $S_{T_i}$  est le regroupement de tous les indices de sensibilités relatifs à la variable  $X_i$  dans une seule mesure, tel que :

$$S_{T_i} = \sum_{p \# i} S_p \quad (2.17)$$

Où  $\#i$  représente tous les ensembles d'indices contenant l'indice  $i$ . Ainsi dans notre exemple précédent avec  $p = 4$ , on aura :

$$S_{T_1} = S_1 + S_{12} + S_{13} + S_{14} + S_{123} + S_{124} + S_{134} + S_{1234}$$

$$S_{T_2} = S_2 + S_{12} + S_{23} + S_{24} + S_{123} + S_{124} + S_{234} + S_{1234}$$

$$S_{T_3} = S_3 + S_{13} + S_{23} + S_{34} + S_{123} + S_{134} + S_{234} + S_{1234}$$

$$S_{T_4} = S_4 + S_{14} + S_{24} + S_{34} + S_{124} + S_{134} + S_{234} + S_{1234}$$

Cette nouvelle réécriture des indices de sensibilité nous permet d'avoir une plus grande vision de l'impact d'une variable d'entrée  $X_i$  sur notre variable de sortie  $Y$ .

### 2.1.3.1 Méthodes d'estimation

Dans le cas des fonctions non linéaires, lorsque la forme de la fonction n'est pas complexe et que les densités de probabilité des variables d'entrées de la fonction sont connues et relativement simples, on peut calculer les indices de sensibilité globale.

Cette situation est peu fréquente dans la réalité, c'est pourquoi il existe d'autres méthodes d'estimation. On utilise des méthodes d'échantillonnage stochastique au niveau des variables

d'entrées pour pouvoir calculer les différentes espérances conditionnelles de la fonction par rapport à ces variables d'entrées. Puis par après on estime les différentes variances (2.15) et indices de sensibilité (2.16).

Une des méthodes les plus couramment employées est une estimation par Monte Carlo ou méthode de Sobol. Il existe aussi cependant d'autres méthodes telles que la méthode FAST (Fourier Amplitude Sensitivity Test) qui utilise la transformée de Fourier multidimensionnelle de  $f$  pour obtenir une décomposition de la variance de  $Y$  ou encore la méthode de McKay qui se base sur un échantillonnage par hypercube latin répliqué (Julien Jaques [12]). Cependant, seule la méthode de Sobol permet le calcul des indices de sensibilité relevant des interactions et est celle qui est la plus rapide à implémenter.

**Exemple :** Nous présentons un exemple d'application tiré de Julien Jaques [6] pour la méthode de Sobol.

Soit le modèle représenté par l'équation :

$$Y = \sin(X_1) + 7\sin^2(X_2) + \frac{X_3^4}{10}\sin(X_1)$$

où  $X_i \sim U[-\pi, \pi]$  avec  $i = 1, 2, 3$ .

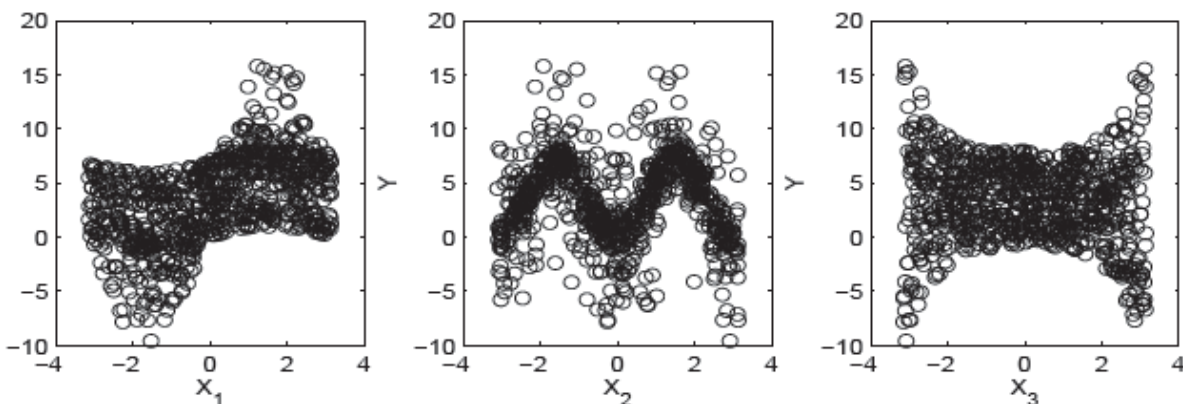


Figure 2-1 Représentation du modèle selon les 3 variables

Les calculs d'indices de sensibilité ont été effectués par l'auteur sous Matlab où il a calculé les indices de sensibilité 20 fois pour des échantillons de taille : 100, 1000 et 10000. On présente dans le tableau ci-dessous la moyenne des indices calculés.

Tableau 2-4 Indices de sensibilité avec la méthode de Sobol

N = 100				N = 1000				N = 10000			
S1	S2	S3	S13	S1	S2	S3	S13	S1	S2	S3	S13
0,318	0,456	0	0,226	0,316	0,449	0	0,235	0,310	0,441	0	0,249

On constate que l'augmentation de la taille de l'échantillon permet de mieux calibrer le modèle et déterminer l'influence de l'interaction ( $X_1, X_3$ ) d'une manière plus précise. De la même manière on constate que l'écart type des estimations d'indices (tableau ci-dessous) diminue avec l'augmentation de la taille des échantillons. Ainsi, si la méthode de Sobol est plus efficace parce qu'elle permet de déterminer les effets d'interactions, elle demande aussi une taille d'échantillon grande pour être plus précise.

Tableau 2-5 Écart types d'estimation de la méthode de Sobol

N = 100			N = 1000			N = 10000		
S1	S2	S3	S1	S2	S3	S1	S2	S3
0,157	0,122	0,150	0,050	0,057	0,073	0,011	0,014	0,018

## 2.2 Méthode d'analyse de sensibilité : forme ou fonction du modèle inconnue

Il existe des cas où la forme du modèle n'est pas connue. Il n'existe pas une fonction pouvant nous permettre de modéliser le phénomène ou du moins on ne connaît pas cette fonction. Dans ces conditions, utiliser les indices introduits plus haut devient impossible dans l'étude de la variance des variables. Il existe cependant des méthodes nous permettant d'analyser dans ce cas-ci l'influence des variables explicatives sur la variable réponse. Nous les aborderons dans les lignes qui suivent.

### 2.2.1 Planification d'expérience

La planification d'expérience est une technique d'analyse de sensibilité globale relative à l'étude de l'influence des variables intrants  $X$  d'un processus sur la variable de sortie  $Y$ . Elle s'applique à tout processus de collecte d'information où il y a présence de variabilité sous le contrôle de

l'expérimentateur. Elle est très souvent utilisée en contrôle de qualité pour étudier et optimiser les processus de production mais son champ d'application s'étend aussi bien au domaine des sciences sociales et humaines. S'il existe plusieurs outils statistiques pour étudier les relations intrants/extrants lorsque leurs domaines de définition sont continus, la planification d'expérience est bien souvent utilisée lorsque les variables changent d'une manière discrète et souvent on fait l'analyse pour des variables ayant un petit nombre de modalités. La planification d'expérience se concentre sur un processus tel que décrit par la figure ci-dessous :

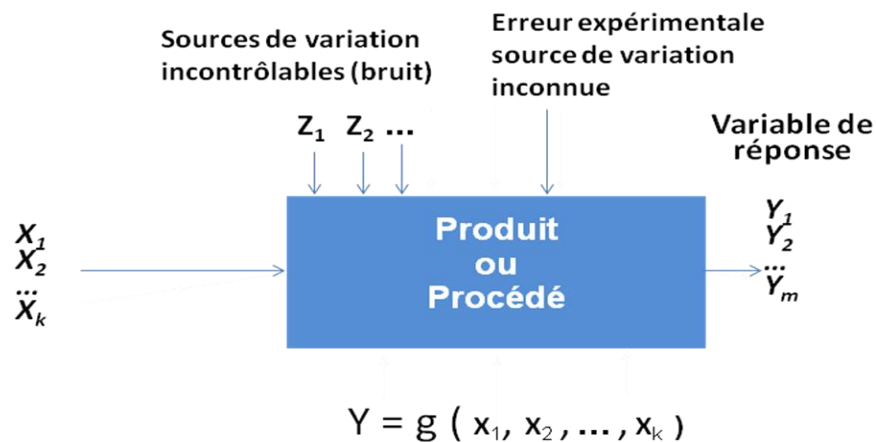


Figure 2-2 Exemple de processus pour la planification d'expérience

A cet égard dans la figure 2.3, la fonction  $g$  de notre processus est inconnue et on fait appel à la planification d'expérience afin d'en développer une approximation. Dans les étapes du processus d'expérimentation, la planification d'expérience regroupe un ensemble d'étapes :

1. définir le processus, la problématique et les objectifs.
2. choisir les variables réponses( $Y$ ) à mesurer.
3. choisir les variables intrants  $X$  et leurs espaces de variations.
4. choisir et comparer les plans expérimentaux.
5. préparer l'expérience (set up).
6. conduire l'expérience.
7. analyser statistiquement les résultats.
8. agir avec les conclusions de l'analyse.

La planification d'expérience permet d'analyser tous les effets des variables intrants du processus, que ce soit leurs effets principaux ou leurs interactions les uns avec les autres, au même titre que les effets de premier ordre et les effets d'ordres supérieurs précédemment discutés pour les indices de sensibilité. Il faut noter cependant que ce sont les étapes après la planification et notamment l'analyse statistique, qui nous permettront d'identifier les effets de chaque variable intrant. Le plan d'expérience en lui-même n'est que le schéma directeur nous permettant d'avoir théoriquement l'ensemble des tests à effectuer pour évaluer le lien entre intrants et extrants du processus. Il existe plusieurs types de plan d'expérience pouvant nous permettre d'une manière plus ou moins complète d'analyser les effets des variables intrants du processus. Ce sont entre autres plans, lorsque les variables du modèle comportent deux modalités et que l'on a  $k$  (entier) variables, les plans : complets, Plackett-Burman, Carré-latin, etc. On utilisera souvent les plans complets pour les expériences puisqu'ils analysent tous les effets. Cependant la limitation par les coûts d'expérimentation ou la répétition d'un effet déjà étudié peuvent nous amener entre autres raisons à diminuer les effets.

### ***Exemple***

*Tiré de Box, G. E. P., Draper, N.R. (1987). Empirical Model-Building and Response Surfaces, John Wiley & Sons. p. 115.*

On cherche à identifier l'influence de six variables à deux modalités lors d'un processus de coloration sur une variable représentant la résistance de cette même coloration. Le tableau ci-dessous nous donne la description des six variables ainsi que les valeurs de leurs différentes modalités.

Tableau 2-6 Variables de l'expérience

Variables	Description	Modalités
X1	Indice de polysulfures	[6,7]
X2	Le taux de reflux lors du processus	[150, 170]
X3	La quantité de polysulfures en moles	[1,8 ; 2,4.]
X4	Le temps en minutes	[24, 36]
X5	Le volume de solvant utilisé en centimètres cubes	[30, 42]
X6	La température lors de la coloration en degrés.	[120, 130]

Ainsi, afin de pouvoir effectuer l'expérience d'une manière optimale et de tester tous les effets pour ensuite calculer les indices de sensibilité totaux tels qu'introduits à l'équation (2.17), on a établi un plan d'expérience complet avec 64 essais ( $2^6$ ). La matrice du tableau 2-7 ci-dessous nous montre ce plan d'expérience et le tableau 2-8 constitue le tableau d'analyse de la variance après expérimentation.

Tableau 2-7 Matrice du Plan d'expérience

	X1polysul	X2reflux	X3moles	X4time	X5solvent	X6temp	Y1_ST
1	6	150	1,8	24	30	120	3,4
2	7	150	1,8	24	30	120	9,7
3	6	170	1,8	24	30	120	7,4
4	7	170	1,8	24	30	120	10,6
5	6	150	2,4	24	30	120	6,5
6	7	150	2,4	24	30	120	7,9
7	6	170	2,4	24	30	120	10,3
8	7	170	2,4	24	30	120	9,5
9	6	150	1,8	36	30	120	14,3
10	7	150	1,8	36	30	120	10,5
11	6	170	1,8	36	30	120	7,8
12	7	170	1,8	36	30	120	17,2
13	6	150	2,4	36	30	120	9,4
14	7	150	2,4	36	30	120	12,1
15	6	170	2,4	36	30	120	9,5
16	7	170	2,4	36	30	120	15,8
17	6	150	1,8	24	42	120	8,3
18	7	150	1,8	24	42	120	8
19	6	170	1,8	24	42	120	7,9
20	7	170	1,8	24	42	120	10,7
21	6	150	2,4	24	42	120	7,2
22	7	150	2,4	24	42	120	7,2
23	6	170	2,4	24	42	120	7,9
24	7	170	2,4	24	42	120	10,2
25	6	150	1,8	36	42	120	10,3
26	7	150	1,8	36	42	120	9,9
27	6	170	1,8	36	42	120	7,4
28	7	170	1,8	36	42	120	10,5
29	6	150	2,4	36	42	120	9,6
30	7	150	2,4	36	42	120	15,1
31	6	170	2,4	36	42	120	8,7
32	7	170	2,4	36	42	120	12,1
33	6	150	1,8	24	30	130	12,6
34	7	150	1,8	24	30	130	10,5
35	6	170	1,8	24	30	130	11,3
36	7	170	1,8	24	30	130	10,6
37	6	150	2,4	24	30	130	8,1
38	7	150	2,4	24	30	130	12,5
39	6	170	2,4	24	30	130	11,1
40	7	170	2,4	24	30	130	12,9
41	6	150	1,8	36	30	130	14,6
42	7	150	1,8	36	30	130	12,7
43	6	170	1,8	36	30	130	10,8
44	7	170	1,8	36	30	130	17,1
45	6	150	2,4	36	30	130	13,6
46	7	150	2,4	36	30	130	14,6
47	6	170	2,4	36	30	130	13,3
48	7	170	2,4	36	30	130	14,4
49	6	150	1,8	24	42	130	11
50	7	150	1,8	24	42	130	12,5
51	6	170	1,8	24	42	130	8,9
52	7	170	1,8	24	42	130	13,1
53	6	150	2,4	24	42	130	7,6
54	7	150	2,4	24	42	130	8,6
55	6	170	2,4	24	42	130	11,8
56	7	170	2,4	24	42	130	12,4
57	6	150	1,8	36	42	130	13,4
58	7	150	1,8	36	42	130	14,6
59	6	170	1,8	36	42	130	14,9
60	7	170	1,8	36	42	130	11,8
61	6	150	2,4	36	42	130	15,6
62	7	150	2,4	36	42	130	12,8
63	6	170	2,4	36	42	130	13,5
64	7	170	2,4	36	42	130	15,8



Tableau 2-8 Analyse de la variance

Numéro	effet	SS	df	MS
1	X1	48,83	1	48,83
2	X2	7,91	1	7,91
3	X3	0,17	1	0,17
4	X4	142,50	1	142,50
5	X5	2,76	1	2,76
6	X6	115,83	1	115,83
7	x1x2	12,69	1	12,69
8	x1x3	0,32	1	0,32
9	x1x4	0,35	1	0,35
10	x1x5	2,76	1	2,76
11	x1x6	10,81	1	10,81
12	x2x3	5,70	1	5,70
13	x2x4	11,82	1	11,82
14	x2x5	1,79	1	1,79
15	x2x6	0,51	1	0,51
16	x3x4	2,60	1	2,60
17	x3x5	0,10	1	0,10
18	x3x6	0,74	1	0,74
19	x4x5	1,59	1	1,59
20	x4x6	0,00	1	0,00
21	x5x6	1,13	1	1,13
22	x1x2x3	6,83	1	6,83
23	x1x2x4	10,64	1	10,64
24	x1x2x5	1,18	1	1,18
25	x1x2x6	1,03	1	1,03
26	x1x3x4	2,60	1	2,60
27	x1x3x5	0,07	1	0,07
28	x1x3x6	0,19	1	0,19
29	x1x4x5	1,72	1	1,72
30	x1x4x6	5,01	1	5,01
31	x1x5x6	0,17	1	0,17
32	x2x3x4	2,60	1	2,60
33	x2x3x5	0,24	1	0,24
34	x2x3x6	2,00	1	2,00
35	x2x4x5	1,79	1	1,79
36	x2x4x6	1,41	1	1,41
37	x2x5x6	5,23	1	5,23
38	x3x4x5	8,19	1	8,19
39	x3x4x6	0,47	1	0,47
40	x3x5x6	0,83	1	0,83
41	x4x5x6	6,31	1	6,31
42	x1x2x3x4	0,74	1	0,74
43	x1x2x3x5	6,70	1	6,70
44	x1x2x3x6	1,35	1	1,35
45	x1x2x4x5	21,51	1	21,51
46	x1x2x4x6	1,59	1	1,59
47	x1x2x5x6	0,01	1	0,01
48	x1x3x4x5	1,86	1	1,86
49	x1x3x4x6	9,38	1	9,38
50	x1x3x5x6	6,83	1	6,83
51	x1x4x5x6	3,56	1	3,56
52	x2x3x4x5	0,93	1	0,93
53	x2x3x4x6	2,68	1	2,68
54	x2x3x5x6	2,44	1	2,44
55	x2x4x5x6	0,02	1	0,02
56	x3x4x5x6	1,24	1	1,24
57	x1x2x3x4x5	7,77	1	7,77
58	x1x2x3x4x6	8,78	1	8,78
59	x1x2x3x5x6	4,05	1	4,05
60	x1x2x4x5x6	3,56	1	3,56
61	x1x3x4x5x6	8,05	1	8,05
62	x2x3x4x5x6	0,32	1	0,32
63	x1x2x3x4x5x6	1,86	1	1,86

Ainsi, nous introduisons une innovation par rapport au calcul des effets dans les planifications d'expériences traditionnelles en calculant non seulement l'apport dans la variance de Y de chaque type d'effet (effets de premier ordre et effets d'interactions) tel que donné par le tableau 2-8, mais nous calculons également la sensibilité totale de Y par rapport à chaque variable d'entrée en calculant à la manière de Saltelli (équation (2.17)) la sensibilité totale dans la variance de la variable Y de chacune des six variables d'entrée. On obtient donc :

Tableau 2-9 Sensibilité totale appliquée à la planification d'expérience

Part de la variance reliée à :						
X1	X2	X3	X4	X5	X6	Total
192,78	137,66	98,61	267,44	106,57	207,39	1010,42
Part de la variance en pourcentage du total reliée à :						
$S_{T_1}$	$S_{T_2}$	$S_{T_3}$	$S_{T_4}$	$S_{T_5}$	$S_{T_6}$	Totale
0,19	0,14	0,10	0,26	0,11	0,21	1

Il en ressort que les variables se classent par ordre d'influence sur la variable réponse selon leurs sensibilités totales comme suit : X4, X6, X1, X2, X5, X3.

## 2.3 Conclusion

Enfin, nous vous présentons une figure (2-3) tirée de Bertrand Iooss [5] qui à notre sens résume bien l'utilisation qui devrait être faite des différentes méthodes que l'on vient de présenter selon la complexité du modèle que l'on veut étudier et du nombre d'évaluations que l'on peut faire pour l'étudier. Ce dernier étant calculé comme un multiple du nombre « d » de variables d'entrée du modèle. De la même manière, Bertrand Iooss [5] introduit le diagramme à la figure 2-4 ci-dessous comme un outil d'aide à la décision de la méthode d'analyse de sensibilité à adopter selon la linéarité du modèle, sa monotonie et le coût pour réaliser l'analyse.

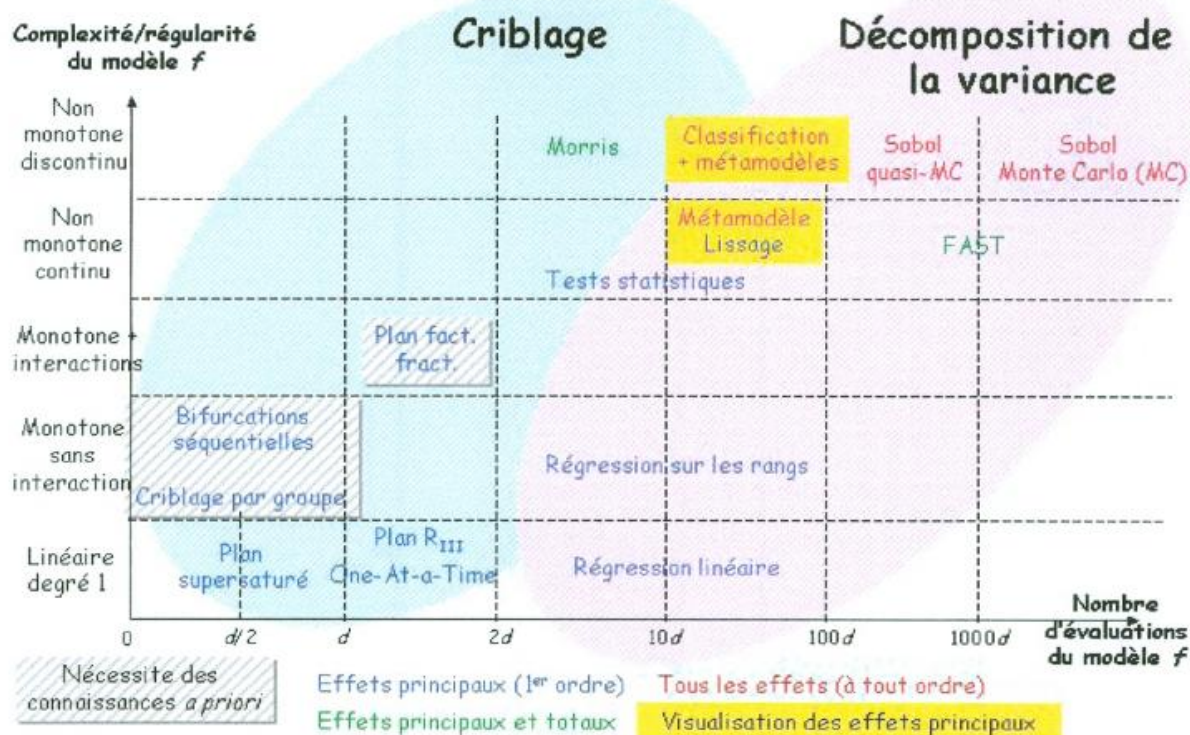


Figure 2-3 Synthèse des méthodes d'analyse de sensibilité

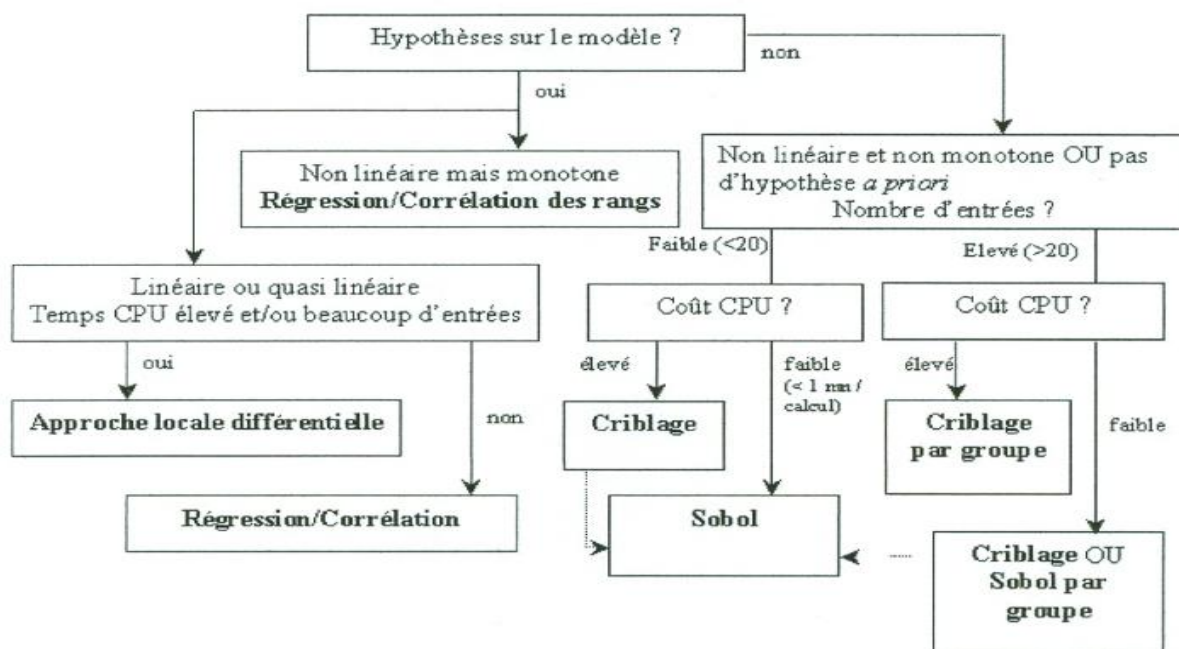


Figure 2-4 Diagramme de décision pour le choix de la méthode d'analyse de sensibilité

## **CHAPITRE 3    MODÉLISATION DU MARCHÉ DES MÉTAUX**

### **3.1 Introduction**

Dans le cadre de notre mémoire nous avons décidé de construire un modèle de prédiction d'un indice reflétant les prix sur les marchés des métaux. Cet indice publié par le « Bureau of Labor Statistics » (BLS [19]) du gouvernement américain, s'appelle WPU 10 « Metals and Metal products ». On va reconstruire synthétiquement cet indice et on testera par la suite la robustesse de l'indice synthétique et sa capacité à prédire le WPU10 d'une manière satisfaisante en termes d'erreur de prédiction. On cherche à voir dans quelle mesure l'évolution dans le temps de l'indice synthétique construit grâce à l'analyse de sensibilité faite sur le WPU10 est robuste face à la variabilité des sous-composants du WPU10.

Dans les paragraphes qui suivent nous donnons quelques informations sur l'indice et sa composition ainsi que l'institution qui le publie. De même, nous faisons un bref rappel des notions auxquelles on fera appel dans le cadre de notre modélisation et nous expliquerons le choix de notre modèle de prévision.

### **3.2 L'indice et l'institution**

#### **3.2.1 Présentation du BLS**

Le BLS [19] est le bureau américain de la statistique du travail. Il est responsable de la publication de la plupart des indices connus et utilisés couramment dans les journaux et dans les institutions économiques. Il publie entre autres indices le fameux indice de l'inflation des prix à la consommation ou CPI (Consumer Price Index) ou les fameux PPI (Producer Price Index) dont le WPU10 fait partie. Ces indices reflètent en quelque sorte l'inflation dans leurs différents domaines et sont donc des agrégats de l'évolution moyenne des prix.

#### **3.2.2 Introduction à la notion du PPI**

Le Producer Price Index (PPI) représente un indice global des prix des produits. Le WPU10 qui reflète le marché des métaux fait partie des PPI, et il est inclus dans ce qu'on appelle les commodités industrielles, codé sous l'indice WPU03THRU15 (se lit WPU 3 à 15, soit donc

WPU 03, WPU 04,..., WPU 15). Ils représentent à eux 13 les différentes matières qui rentrent dans le processus de production industrielle.

Tableau 3-1 Décomposition des PPI

Classification par code des PPIs (WPU)	Description	Poids (%)
<b>03Thru15</b>	<b>Industrial Commodities</b>	<b>100,0000</b>
3	Textile products and apparel	2,3295
4	Hides, skins, leather, and related products	0,1838
5	Fuels and related products and power	21,1646
6	Chemicals and allied products	12,5831
7	Rubber and plastic products	4,3090
8	Lumber and wood products	2,0809
9	Pulp, paper, and allied products	9,2626
<b>10</b>	<b>Metals and Metal products</b>	<b>11,5603</b>
11	Machinery and equipment	14,1808
12	Furniture and household durables	2,5828
13	Nonmetallic mineral products	3,0671
14	Transportation equipment	12,2001
15	Miscellaneous products	4,4964

On peut noter dans le tableau 3.1 ci-dessus que les produits regroupés sous la dénomination commodités industrielles incluent un large éventail allant de matières organiques à des matières minérales. Cependant, il faut noter que ces matières sont conçues comme des produits bruts ou semi transformés qui entrent dans la production d'autres produits finis.

Dans le tableau ci-dessus le cas qui nous intéresse est celui du WPU10 (surligné en gris) dont la sous-composition se présente comme suit dans le tableau 3.2 :

Tableau 3-2 Décomposition du WPU10

Rang dans l'échelle de classification de parenté	Classification par code des PPIs (WPU)	Description	Poids (%)
1	10	Metals & Metal Products	100
2	101	Iron and steel	26,521
3	101211	Carbon steel scrap	3,109
3	1015	Foundry and forge shop products	4,434
3	10170301	Hot rolled sheet and strip, incl. tin mill	3,68
3	10170401	Hot rolled bars, plates, and structural shape	3,771
3	10170601	Steel pipe and tube	3,72
3	10170701	Cold rolled sheets and strip	3,2
2	102	Nonferrous metals	22,567
3	1021	Nonferrous metal ores	2,834
5	102201	Primary nonferrous metals	3,037
3	1023	Nonferrous scrap	2,059
5	102501	Aluminum mill shapes	3,251
5	102502	Copper and brass mill shapes	1,906
3	1026	Nonferrous wire and cable	3,965
3	1028	Nonferrous foundry shop products	2,732
3	1031	Metal cans and can components	2,487
2	104	Hardware	3,465
3	1071	Metal doors, sash, and trim	2,304
3	1073	Sheet metal products	3,384
3	1074	Struct., arch., pre-eng. metal products	8,062
2	108	Miscellaneous metal products	19,254
2	109	Metal treatment services	3,904

Tout comme dans l'exemple de l'indice *WPU03THRU15* on a une décomposition de l'indice *WPU10* fournie par le BLS. Dans le cas du tableau 3.2, les sous-composants ont une configuration d'ascendance. Ainsi, on peut remarquer que tous les codes d'indices dans le tableau commencent par 10, ce qui signifie que dans la classification du BLS ils sont sous *WPU10*.

De même on peut remarquer que le *WPU101*, *WPU102* et tous les autres indices surlignés en gris sont des indices mères de deuxième niveau bien que certains n'ont pas de sous-composants. Cependant dans le cas du *WPU101* et *WPU102* leurs sous-composants incluent respectivement dans leurs codifications les chiffres 101 (ex : *WPU101211*) et 102 (*WPU1021*).

Par ailleurs, dans le tableau 3.2, la liste des sous-composants du *WPU10* n'est pas exhaustive. Nous avons suivi la décomposition de l'indice donnée par le BLS jusqu'à obtenir des poids de sous-composants supérieurs ou égal à 2% du poids total de *WPU10*. On a gardé

exceptionnellement les indices dont le poids est inférieur mais proche de 2%, ce que nous pensons être significativement important pour nous permettre de faire une modélisation d'un indice synthétique *WPU10* par la suite. Nous avons toutefois omis tous les sous-composants avec un poids significativement inférieur à 2%.

Les données sur ces différents indices fournies par le BLS se présentent sous la forme :

Tableau 3-3 Données du WPU03THRU15 selon le BLS

**Series Id: WPU03THRU15**

**Not Seasonally Adjusted**

**Item:** Industrial commodities

**Base** 198200

**Date:**

Year	Jan	Feb	Mar	Apr.	May	Jun	July	Aug	Sep	Oct.	Nov.	Déc.	Annual
2000	130,0	131,5	132,6	132,2	133,0	135,9	135,9	135,2	137,3	137,8	137,3	138,5	134,8
2001	142,9	139,7	137,7	138,2	138,6	137,1	134,5	134,3	134,3	131,1	130,9	129,1	135,7
2002	129,4	129,1	130,5	132,4	132,3	132,4	132,6	132,8	133,7	134,8	134,7	134,4	132,4
2003	136,7	139,3	143,6	138,2	137,8	139,2	139,1	139,1	139,1	139,2	138,8	139,5	139,1
2004	142,2	142,8	143,3	144,8	146,5	147,3	148,2	149,3	149,1	151,8	153,5	152,0	147,6
2005	152,7	153,6	155,6	157,2	156,3	156,6	159,1	160,8	166,0	170,6	167,6	166,5	160,2
2006	168,3	165,7	166,3	168,8	170,6	170,6	171,3	172,4	169,2	165,4	168,0	168,9	168,8
2007	166,8	169,1	171,6	173,9	176,0	176,4	177,9	174,9	175,6	176,9	181,8	180,7	175,1
2008	182,8	184,6	190,2	193,8	200,0	204,0	209,5	202,4	200,1	189,3	178,4	172,3	192,3
2009	172,6	170,8	169,5	170,3	172,0	175,5	174,6	177,7	176,9	177,8	180,1	180,4	174,8
2010	184,6	183,6	185,7	186,9	187,4	187,2	186,7	187,5	186,8				

Tableau 3-4 Données du WPU10 selon le BLS

**Series Id: WPU10**

**Not Seasonally Adjusted**

**Item:** Metals and metal products

**Base** 198200

**Date:**

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Annual
2000	128,3	128,8	128,7	128,6	128,2	127,9	128,0	128,0	128,4	127,9	126,8	126,9	128,1
2001	126,9	126,6	126,6	126,0	126,1	125,8	125,5	125,0	124,9	124,1	123,6	123,6	125,4
2002	123,7	124,0	124,5	125,0	125,6	126,4	126,8	126,6	127,1	127,0	127,3	127,2	125,9
2003	127,6	128,3	128,5	128,2	128,3	128,3	128,4	129,0	129,5	130,2	131,4	133,1	129,2
2004	135,9	140,2	143,9	146,5	147,0	147,3	151,3	154,0	154,7	157,1	158,6	159,0	149,6
2005	160,1	160,5	160,4	161,1	159,4	157,6	157,4	158,4	161,1	161,9	165,0	166,7	160,8
2006	168,6	170,9	172,0	176,9	184,2	184,9	187,5	187,0	187,7	187,3	186,0	186,5	181,6
2007	185,7	187,2	191,1	195,4	196,3	195,9	196,6	195,5	194,5	195,0	194,3	194,1	193,5
2008	197,5	201,8	208,0	217,6	223,4	226,9	231,8	230,9	223,7	209,1	195,9	189,7	213,0
2009	187,0	183,9	181,7	179,9	180,5	181,7	183,5	189,1	192,1	193,3	193,0	196,0	186,8
2010	200,5	200,5	204,5	210,5	210,4	207,4	205,0	206,5	208,2				

Les valeurs mensuelles des indices dans les tableaux ci-dessus sont calculées à partir d'une année de base où l'on a calculé une moyenne de prix des matières premières composant chaque indice que l'on a fixé à 100. Puis, sur une base mensuelle on a recalculé cette moyenne et affecté sa variation à la base indicielle de 100.

Ainsi par exemple, en janvier 1982 si la moyenne des prix des matières premières composant le WPU10 était de 1000 et qu'en février 1982 cette moyenne était de 1010, si l'on a fixé le WPU10 à janvier 1982 à 100 sa valeur en février serait de  $100 \times (1010/1000)$  soit 101.

### **3.3 Rappel de quelques notions utilisées dans la modélisation**

Dans le cadre de ce mémoire on a fait appel à certaines notions en mathématiques et en finance afin de pouvoir effectuer notre modélisation. Dans les lignes qui suivent, nous les introduisons.

#### **3.3.1 Les contrats futurs**

Dans le marché financier il est possible pour un investisseur d'acheter en avance un produit à un prix déterminé grâce à un contrat futur. Ainsi, par exemple au 1<sup>er</sup> juin 2010 un investisseur pourra accepter un contrat futur pour l'achat d'une tonne de cuivre qui lui sera livrée au 1<sup>er</sup> septembre 2010 à un prix de 1100\$ alors que le prix en juin 2010 est de 1000\$ /tonne.

Cela voudrait dire qu'au 1<sup>er</sup> juin 2010 les investisseurs sur le marché financier pensent que le prix d'une tonne de cuivre va croître de 10% sur les trois prochains mois. Cependant comme les prix sont cotés presque instantanément, ils sont sujets à des variations. Celles-ci, même si elles sont minimales à l'intérieur d'une même journée, peuvent être grandes lorsque le temps augmente. Ainsi par exemple le lendemain ou la semaine d'après le prix pourra monter à 1300\$.

#### **Exemple**

Soit un investisseur sur le marché financier qui est intéressé à connaître l'évolution du prix du cuivre sur les douze prochains mois à partir d'un temps initial zéro. Les données qu'il obtient du marché financier pour le prix futur d'une tonne de cuivre sont :



Tableau 3-5 Prix futurs d'un tonne de cuivre

Prix d'une tonne de cuivre													
Mois	0	1	2	3	4	5	6	7	8	9	10	11	12
Prix	1000	900	1100	1000	1050	1120	1150	1090	1100	1050	1100	1150	1200

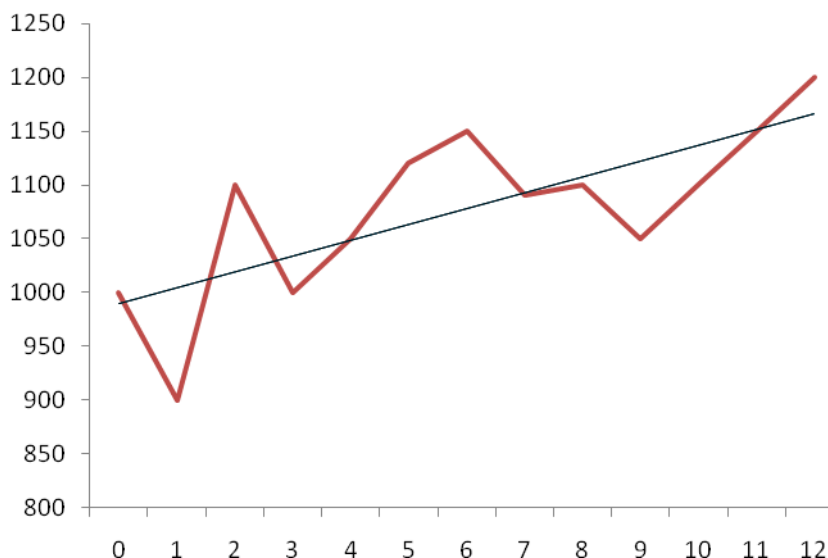


Figure 3-1 Prix futur d'une tonne de cuivre sur douze mois

On peut voir qu'à l'instant  $t = 0$  l'investisseur a une vision sur la tendance future à long terme tel que représenté par la figure ci-dessus. Le concept de contrats futurs est essentiel pour nous dans ce mémoire où l'on cherche à faire une prévision dans le futur sur le marché financier. En effet, il nous permet à un instant donné d'avoir une vision des prix futurs des matières premières qui nous intéresse.

### 3.3.2 Modèle de diffusion et processus stochastique

Il existe plusieurs façons de calculer la valeur future d'un actif financier : régression, modèle autorégressif, etc. Cependant dans le cadre de ce mémoire où l'on cherche à étudier la puissance de l'analyse de sensibilité comme outil d'analyse et de prévision, celui qui nous intéresse est le du modèle de diffusion. Nous allons présenter le modèle de diffusion tel que décrit par François

Pascal [3] qui lui-même le reprend de Black & Scholes et nous allons expliquer par après en quoi il peut nous être utile dans un cadre de prévision des prix futurs des matières premières.

Le cas du modèle de diffusion est proposé ici dans le cadre d'un actif financier pour faire le lien avec les matières premières que nous étudions et qui sont transigées sur les marchés financiers. Ainsi, dans le cadre du marché financier l'actif est transigé d'une manière quasi continue, et le prix de ce dernier est reconnu comme suivant un processus stochastique.

Soit  $S_t$  le prix de l'actif à un temps  $t$  et soit  $S_0$  le prix de l'actif au temps initial 0, tous deux différents de zéro. Posons

$$Z_{0,t} = \frac{S_t - S_0}{S_0}$$

$Z_{0,t}$  correspond au rendement de l'actif sur la période  $[0, t]$ . De même, pour tout temps  $t > 0$ , on a :

$$Z_{t,t+1} = \frac{S_{t+1} - S_t}{S_t}$$

François Pascal [3] démontre que ce processus discret a pour espérance et variance :

$$\begin{aligned} E(Z_{t,t+1}) &= \mu \\ \text{Var}(Z_{t,t+1}) &= \sigma^2 \end{aligned}$$

Et par ailleurs pour tout  $(s, t)$  où  $s > t$ , l'espérance et la variance de  $Z_{t,s}$  sont proportionnelles à  $s-t$ . En particulier pour tout  $t$  :

$$E(Z_{0,t}) = \mu t$$

$$\text{Var}(Z_{0,t}) = \sigma^2 t$$

De même, l'auteur montre que le processus discret  $(Z_{0,t})_{t \geq 0}$  a une distribution normale. Ce qui lui permet de réécrire le rendement de l'actif  $Z_{0,t}$  avec  $U$  une variable aléatoire suivant la loi normale centrée réduite comme :

$$Z_{0,t} = \frac{S_t - S_0}{S_0}$$

$$S_t - S_0 = S_0 Z_{0,t}$$

$$S_t - S_0 = S_0 \mu t + S_0 \sigma \sqrt{t} U$$

$$S_t = S_0 (1 + \mu t + \sigma \sqrt{t} U) \quad (3.1)$$

Il apparaît donc que dans le cas des actifs financiers, la valeur prévue de l'actif selon (3.1) est une somme entre la valeur espérée  $S_0(1 + \mu t)$  de l'actif d'une part, et d'autre part, une partie  $S_0 \sigma \sqrt{t} U$  qui correspond à une variation qui contient une composante stochastique.

### Exemple

Si nous reprenons notre exemple de l'investisseur intéressé par les prix du cuivre du point (3.3.1) sur les contrats futurs, on avait comme information :

Tableau 3-6 Prix futurs du cuivre

Prix d'une tonne de cuivre par mois													
Mois	0	1	2	3	4	5	6	7	8	9	10	11	12
Prix	1000	900	1100	1000	1050	1120	1150	1090	1100	1050	1100	1150	1200

Ces prix dans le cadre de l'équation (3.1) correspondent à la valeur espérée du prix soit à la partie  $S_0(1 + \mu t)$  de l'équation, puisque  $S_0 \sigma \sqrt{t} U$  devient nulle lorsqu'on remplace  $U$  par  $E(U)$ . Le taux  $\mu t$  correspond au taux de croissance cumulé pour chaque mois par rapport au temps initial, il correspond au rendement potentiel sur la période.

Tableau 3-7 Taux de croissance mensuel du prix du cuivre basé sur le prix futur

	Taux de croissance du temps initial au temps t												
	0	1	2	3	4	5	6	7	8	9	10	11	12
Taux de croissance mensuel ( $\mu$ )		-10,0%	5,0%	0,0%	1,3%	2,4%	2,5%	1,3%	1,3%	0,6%	1,0%	1,4%	1,7%

Par ailleurs si le taux de croissance mensuel du prix du cuivre a un écart type mensuel de 5% et que pour la loi normale centrée réduite le quantile 60% est de 0,2533 et le quantile 40% est de -0,2533, l'investisseur grâce à la fonction de répartition de la loi normale peut déterminer des intervalles de confiance pour ses différents taux de croissance en appliquant l'équation (3.1).

Dans le tableau ci-dessous, on calcule les prix selon l'équation (3.1) en remplaçant U par les valeurs de ses quantiles pour déterminer un intervalle de confiance pour les prix:

Tableau 3-8 Exemple d'application du processus de diffusion

Prix d'une tonne de cuivre par mois scénario du quantile 40%													
Mois	0	1	2	3	4	5	6	7	8	9	10	11	12
Prix espéré	1000	900	1100	1000	1050	1120	1150	1090	1100	1050	1100	1150	1200
+													
écart type du taux de croissance mensuel	5%	5%	5%	5%	5%	5%	5%	5%	5%	5%	5%	5%	5%
X													
Racine de t	0,000	1,000	1,414	1,732	2,000	2,236	2,449	2,646	2,828	3,000	3,162	3,317	3,464
Quantile 60%	-0,2533	-0,2533	-0,2533	-0,2533	-0,2533	-0,2533	-0,2533	-0,2533	-0,2533	-0,2533	-0,2533	-0,2533	-0,2533
Prix avec la probabilité du quantile 40%	1000,00	887,33	1082,09	978,06	1024,67	1091,67	1118,97	1056,49	1064,17	1012,00	1059,94	1107,99	1156,12
Prix d'une tonne de cuivre par mois scénario du quantile 60%													
Mois	0	1	2	3	4	5	6	7	8	9	10	11	12
Prix espéré	1000	900	1100	1000	1050	1120	1150	1090	1100	1050	1100	1150	1200
+													
écart type du taux de croissance mensuel	5%	5%	5%	5%	5%	5%	5%	5%	5%	5%	5%	5%	5%
X													
Racine de t	0,000	1,000	1,414	1,732	2,000	2,236	2,449	2,646	2,828	3,000	3,162	3,317	3,464
Quantile 60%	0,2533	0,2533	0,2533	0,2533	0,2533	0,2533	0,2533	0,2533	0,2533	0,2533	0,2533	0,2533	0,2533
Prix avec la probabilité du quantile 60%	1000,00	912,67	1117,91	1021,94	1075,33	1148,33	1181,03	1123,51	1135,83	1088,00	1140,06	1192,01	1243,88

Ce dernier tableau nous permet d'établir la figure ci-dessous de la prévision du prix d'une tonne de cuivre ainsi que ses possibles évolutions dans l'intervalle de confiance défini.

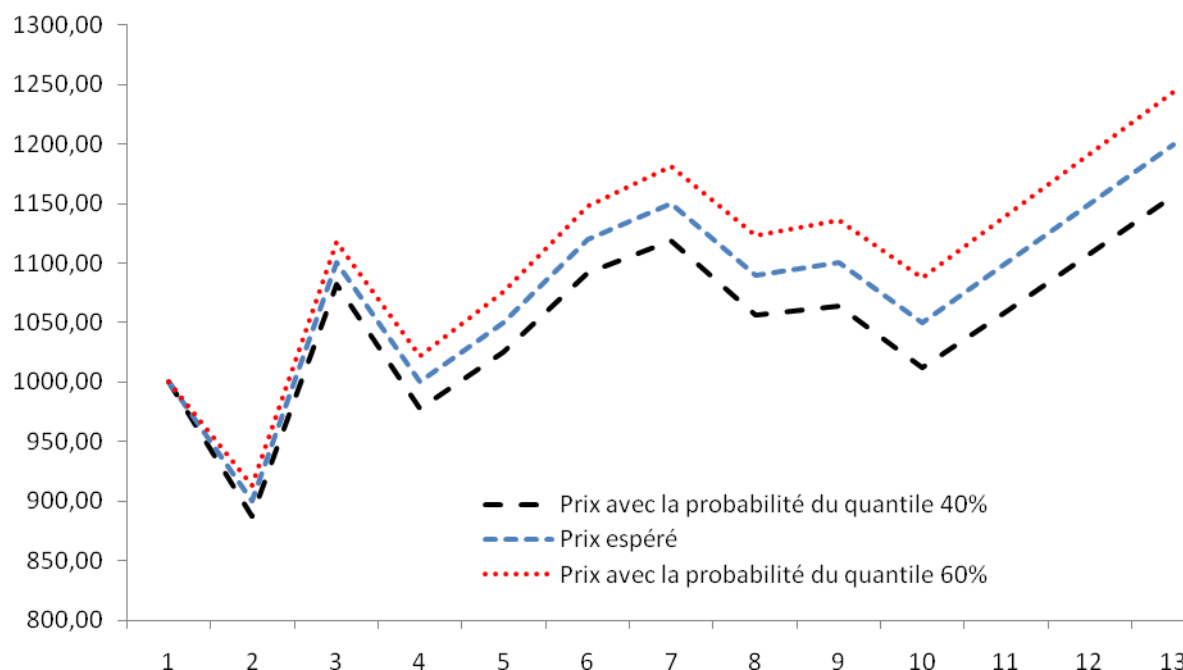


Figure 3-2 Intervalle de confiance du prix futur d'une tonne de cuivre sur 12 mois

Cette approche traduit parfaitement le fait que bien qu'à un temps donné au travers des contrats futurs on sache ce que peut être notre taux de croissance espéré, il existe un risque que ce taux change. Cependant l'équation (3.1) nous apprend que ce changement bien qu'aléatoire (selon la variable aléatoire  $U$ ), est fonction de la variance historique ce qui nous permet d'établir des intervalles de confiance dans lesquels on peut évaluer le prix futur. On calcule ainsi grâce à la variable aléatoire des variations de notre prix futur autour des valeurs espérées.

### 3.3.3 Décomposition de Cholesky : les nombres aléatoires multivariés.

Nous avons présenté dans la section précédente un modèle de prévision faisant appel à des simulations de prix. Nous comptons dans le processus de simulation faire des perturbations sur le prix des matières premières que l'on assimile à nos indices sous-composants du WPU10 pour en effectuer la prévision. Néanmoins, bien qu'aléatoires nous avons établi que ces perturbations suivent une logique sous-jacente. Il existe cependant une autre logique qu'il nous faudra aussi considérer. Jusqu'à présent on a présenté la possibilité de simulation du prix pour un seul actif seulement (le cuivre). Cependant on a établi au tableau 3.2 que l'indice WPU10 est fait de plusieurs sous-indices et par ailleurs la plupart de ces indices sont corrélés entre eux. Par conséquent perturber le prix du cuivre à la hausse avec l'équation (3.1), sans considération du

prix de l'aluminium par exemple lorsque celui-ci est corrélé au cuivre, n'aurait pas de sens dans la mesure où dans le marché réel ces deux actifs sont corrélés. Il nous faut donc trouver une façon de prendre en compte la corrélation entre le prix de ces deux métaux. La décomposition de Cholesky nous permet de répondre à cette question.

**Théorème** *Factorisation ou décomposition de Cholesky*

La **factorisation de Cholesky**, nommée d'après André Louis Cholesky est la règle selon laquelle pour une matrice symétrique semi-définie positive  $A$ , il existe au moins une matrice réelle triangulaire inférieure  $L$  telle que :  $A = LL'$  où  $L'$  est la transposée de  $L$ .

Il est possible d'ajouter une condition telle que les éléments diagonaux de la matrice  $L$  soient tous positifs. Dans ce cas, la factorisation correspondante est unique.

**Exemple d'application de Cholesky**

Soit :

- $G, V, W$  et  $Z$  quatre variables représentant les prix mensuels de quatre matières premières.
- Supposons que  $\Sigma$  est leur matrice de corrélation, tel que  $\Sigma$  s'écrit :

	G	V	W	Z
G	1	0,44	0,33	0,2
V	0,44	1	0,45	0,56
W	0,33	0,45	1	0,34
Z	0,2	0,56	0,34	1

On sait qu'à tout instant  $t$ , si l'on désire perturber les variables  $(X_i, X_j)$  faisant partie de  $(G, V, W, Z)$ ,  $i$  différent de  $j$ , d'une perturbation respectivement de  $\lambda$  et  $\omega$  sur la base de la corrélation  $\rho$  du couple  $(X_i, X_j)$ , il nous faut établir une relation entre la valeur  $\lambda$  utilisée pour perturber  $X_i$  et la valeur  $\omega$  utilisée pour perturber  $X_j$  pour conserver la corrélation  $\rho$  entre le couple  $(X_i, X_j)$  après perturbation. Cela, tout en sachant que les perturbations doivent être aléatoires. Il nous faut donc établir des perturbations aléatoires mais corrélées.

- Soit  $L$  la matrice obtenue après application de la décomposition de Cholesky de la matrice des corrélations  $\Sigma$  tel que  $\Sigma = LL'$ .

- L s'écrit :

	G	V	W	Z
G	1	0	0	0
V	0,44	0,89799777	0	0
W	0,33	0,33942178	0,88084781	0
Z	0,2	0,52561378	0,1085264	0,8197269

- L' s'écrit :

	G	V	W	Z
G	1	0,44	0,33	0,2
V	0	0,89799777	0,33942178	0,52561378
W	0	0	0,88084781	0,1085264
Z	0	0	0	0,8197269

- Soit U un vecteur de nombres aléatoires indépendants, que nous utiliserons indirectement pour perturber les prix des matières premières, tel que U s'écrit :

Variables	U
G	0,66550959
V	0,25288677
W	0,28222891
Z	0,42276092

On peut calculer sur la base de U un vecteur Z de nombre aléatoires corrélés qui serviront à perturber les prix, en posant :

$$Z = LU$$

En effet en calculant Z de cette manière on peut se rendre compte que :

$$\text{Var}(Z) = \text{Var}(LU)$$

$$\text{Var}(Z) = L \text{Var}(U) L'$$

$$\text{Var}(Z) = LL'$$

Car  $\text{Var}(U) = I$

Donc,

$$\text{Var}(Z) = \Sigma$$

or si l'on considère deux variables aléatoires corrélées  $\varepsilon$  et  $\varphi$  distribuées normalement ( $N(0,1)$ ), leur corrélation s'écrit :

$$\begin{aligned}\rho_{\varepsilon,\varphi} &= \frac{\text{cov}(\varepsilon,\varphi)}{\sigma_{\varepsilon}\sigma_{\varphi}} = \text{cov}(\varepsilon,\varphi) \\ &= E[(\varepsilon - E(\varepsilon))(\varphi - E(\varphi))] = E(\varepsilon,\varphi)\end{aligned}$$

Ce qui dans le cas de  $\text{Var}(Z)$  se traduit par :

$$\text{Var}(Z) = E(ZZ^T) = \Sigma$$

La corrélation des variables aléatoires générées sous  $Z$  est donc bien notre  $\Sigma$  initial.

En utilisant  $U$  et  $L$  tel qu'introduit ci-dessus, on peut calculer  $Z$ , tel que  $Z$  s'écrit :

Variables	Z
G	0,66550959
V	0,51991598
W	0,55405416
Z	0,64320048

Les éléments de  $Z$  pourront être utilisés pour perturber la valeur de la variable à laquelle ils font référence sans se soucier du fait que la perturbation soit conforme à la corrélation que chacune des variables a avec les autres.

Par la suite, dans ce mémoire, tout comme dans le cas des contrats futurs, nous ferons appel à ce concept dans notre modèle de prévision.

### 3.4 Le processus de modélisation

Nous nous intéressons à l'indice WPU10 publié par le BLS. Cet indice est composé d'indices sous-composants qui impactent ses variations. Nous voulons utiliser les outils que nous avons au moyen de l'analyse de sensibilité pour pouvoir étudier la variation de cet indice.

Nous présentons un diagramme de notre processus de modélisation que nous détaillons par la suite :



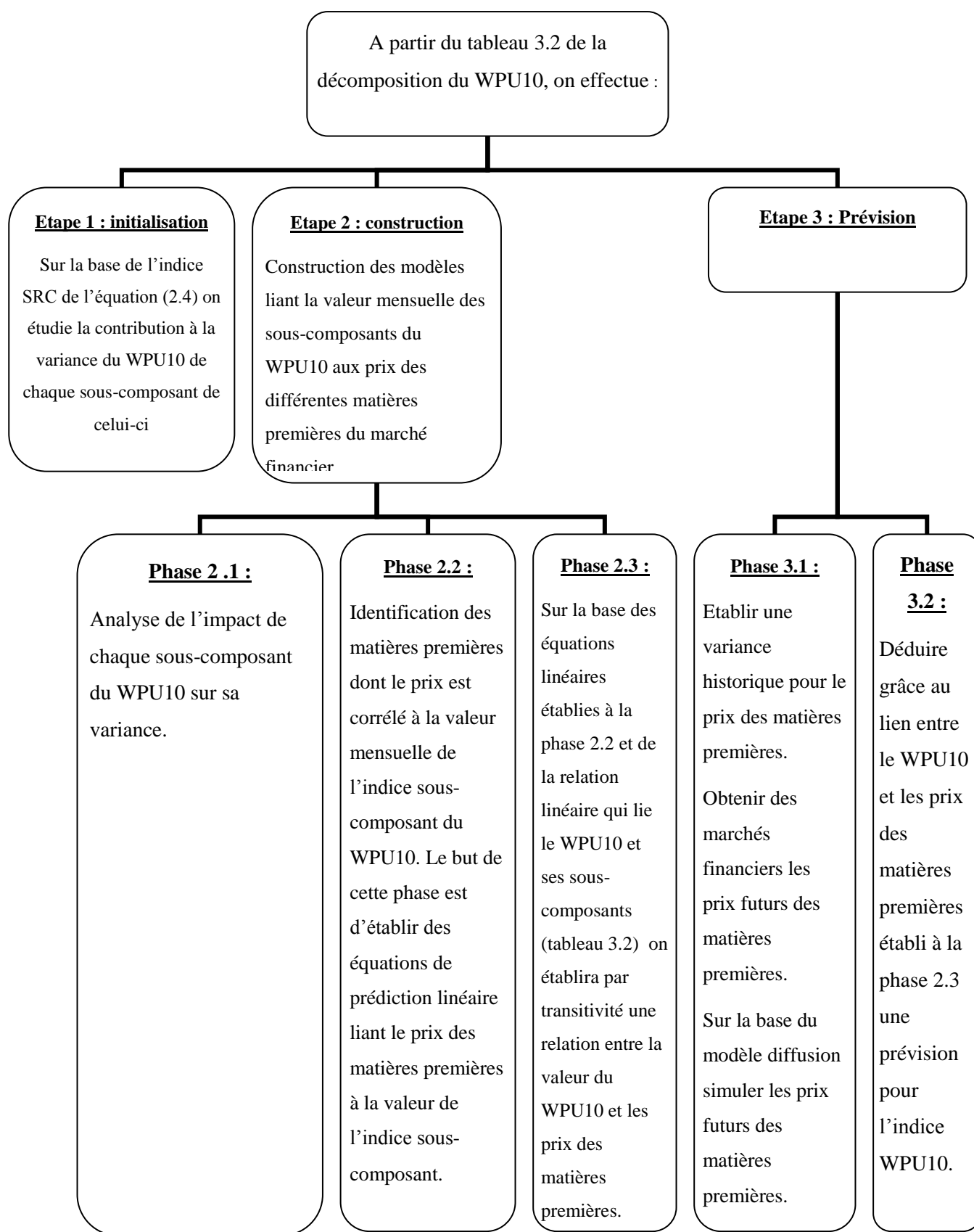


Figure 3-3 Diagramme du processus de modélisation

Notre processus de modélisation se décompose en 3 étapes :

- **Étape 1 :** Comme nous cherchons à faire une prévision du WPU10 comme application de l'analyse de sensibilité, il nous faut comprendre la variance du WPU10 et de ses sous-composants. En l'occurrence, on voudra connaître quels indices influent la variance du WPU10 et quels indices ne l'influent pas. Prédire le WPU10 consiste avant tout à calculer sa variance pour prédire sa valeur. On utilisera pour cela les indices de sensibilité introduits au chapitre 2.
- **Étape 2 :** Cette étape se découpe en trois phases.

**Phase 2.1 :** on calcule la part de variance de chaque sous-composant dans la variance totale du WPU10 et on classe les indices en deux groupes : les sous-composants ayant un impact important et ceux ayant un impact négligeable. Puis pour ceux ayant un impact important, on cherchera à identifier des variables avec lesquelles ils sont corrélés.

**Phase 2.2 :** on utilise les indices sous-composants ayant un impact important sur la variance du WPU10 et pour chacun d'eux on essaie de comprendre leurs comportements. Pour cela on associe leurs comportements à celui de variables que l'on connaît mieux et pour lesquelles on a plus d'informations. Ainsi par exemple, prédire l'évolution de l'indice WPU1031 sous-composant du WPU10 au nom de « Metal cans and products » nous serait difficile. Par contre, si son comportement est corrélé à celui du prix de l'aluminium par exemple, l'exercice devient plus facile puisque l'on dispose de l'information d'une manière continue provenant des marchés financiers sur ce prix. Le but de cette phase est donc pour chaque indice sous-composant du WPU10 d'établir une équation d'une fonction linéaire reliant la valeur du sous-indice au prix d'une matière première. Cette équation est du type :

$$S_{i,t} = G_i(X_t) \quad (3.2)$$

avec,  $S_{i,t}$  la valeur de l'indice sous-composant  $i$  au temps  $t$ ,  $X_t$  le prix au temps  $t$  de la matière première avec laquelle il est corrélé et  $G_i$  la fonction régissant cette relation.

**Phase 2.3 :** Cette phase a pour but de permettre d'associer au moyen d'une équation linéaire les différents prix des matières premières à la valeur de l'indice WPU10. Pour ce faire on raisonnera d'une manière transitive.

On sait d'après le tableau 3.2 que pour tout  $Y_t$  valeur de l'indice WPU10 au temps  $t$ , il existe des  $W_i$  poids des indices  $i$  sous-composants du WPU10, des  $S_{i,t}$  valeurs des indices  $i$  sous-composants du WPU10 au temps  $t$ , et  $n$  un entier naturel correspondant au nombre de sous-composants de WPU10, tel que :

$$Y_t = \sum_{i=1}^n W_i S_{i,t} \quad (3.3)$$

Par ailleurs, on sait aussi d'après l'équation (3.2) que l'on peut avoir une relation entre le prix des matières premières et les sous-composants du WPU10. Il nous est donc possible d'avoir une relation linéaire entre la valeur au temps  $t$  de l'indice WPU10 et le prix affiché sur les marchés financiers de certaines matières premières que l'on désignera comme indicateurs avancés. Cette relation sera du type :

$$Y_t = \sum_{i=1}^n W_i S_{i,t} = \sum_{i=1}^n W_i G_i(X_t) \quad (3.4)$$

On aura donc une nouvelle variable  $Y$  qui nous permettra d'avoir une valeur approchée du WPU10. On désignera cette nouvelle variable comme indice synthétique répliquant le comportement de WPU10.

- **Étape 3 :** après identification des indicateurs avancés des sous-composants du WPU10 et la construction de notre indice synthétique, on cherchera à construire un modèle de prévision à proprement parler. Ce modèle s'exécutera selon deux phases :

**Phase 3.1 :** Prévision du prix des matières premières grâce à l'application de l'équation (3.1).

**Phase 3.2 :** Prévision du WPU10 grâce à l'application de l'équation (3.4).

Dans les prochains paragraphes, on détaillera le modèle de prévision.

## 3.5 Le modèle de prévision

Comme énoncé plus haut notre modèle de prévision est un modèle basé sur une structure à rebours. On effectuera la prévision du prix des matières premières corrélés aux indices sous composants du WPU10 dans une première partie pour ensuite reconstruire synthétiquement au moyen de l'équation (3.4) la valeur du WPU10.

### 3.5.1 Structure du modèle de prévision

Afin de visualiser notre modèle de prévision nous utilisons un diagramme de processus schématisant notre modèle puis nous expliquerons en détail les caractéristiques de l'algorithme qui nous a servi à programmer ce modèle.

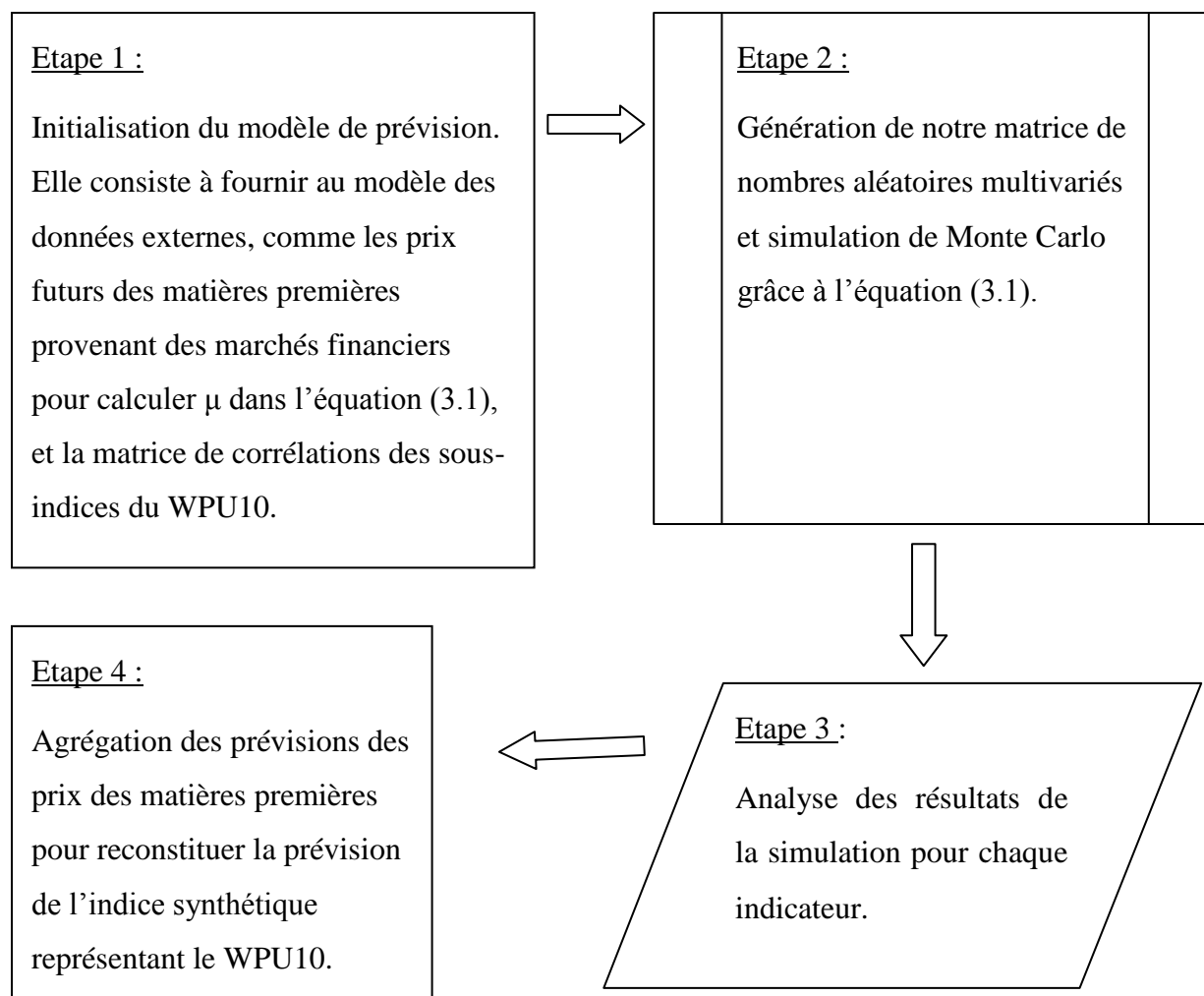


Figure 3-4 Schématisation du processus de prévision

### Description des étapes

On a établi un modèle de diffusion (cf. équation (3.1)) du type :

$$S_t = S_0 + S_0\mu t + S_0\sigma\sqrt{t}U$$

avec

$S_0$  : le prix de l'objet d'intérêt aujourd'hui (valeur de l'indice, prix d'une tonne de cuivre, etc.)

$\mu$  : taux de croissance espéré de S

$\sigma$  : écart type de la croissance S

U : représente un processus stochastique, que nous avons supposé distribué selon une loi  $N(0,1)$

ce qui nous permet de prédire la valeur de S à tout moment t, et ainsi obtenir grâce au processus stochastique U par le biais de notre matrice de nombres aléatoires multivariés plusieurs scénarios futurs. A la fin de la simulation on a un résultat final résumé par le graphique ci-dessous :

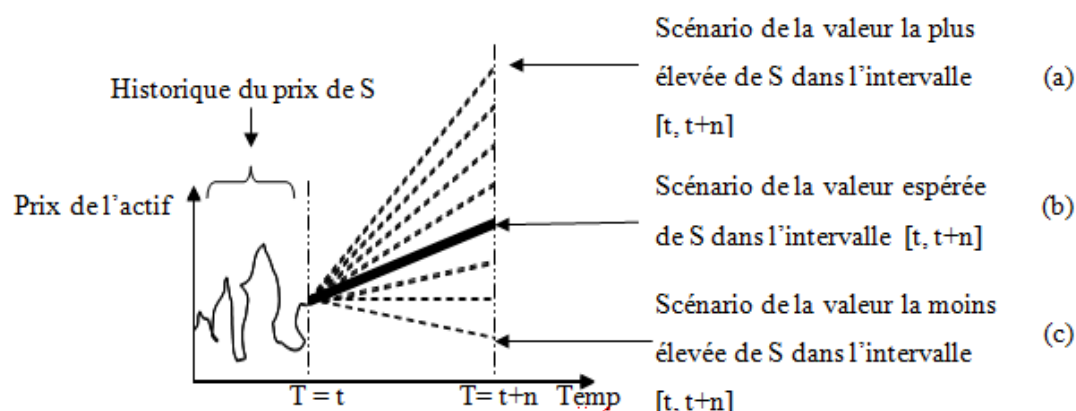


Figure 3-5 Schématisation des résultats de la prévision sur un intervalle de temps  $[t, t+n]$

Pour atteindre un résultat tel que représenté par la figure ci-dessus, nous suivons plusieurs étapes tel qu'indiqué par le diagramme de processus de la figure 3.1.

**Étape 1 :** Dans la première étape de notre prévision, nous allons pourvoir notre modèle de données essentielles à la simulation, dont entre autres les valeurs espérées futures de S sur

l'intervalle de prévision  $[t, t+n]$  qui nous intéresse. Dans cette étape nous irons chercher de sources externes (bases de données des prix sur les marchés financiers tels que Bloomberg, Reuters, etc.) les données de corrélation pour les métaux (calculées grâce aux valeurs historiques de  $S$ ) qui nous intéressent, ainsi que leurs prix futurs tel qu'anticipés par le marché financier et leurs écarts types.

Les coefficients de corrélation serviront essentiellement à générer des nombres aléatoires multivariés grâce à la factorisation de Cholesky, ce qui nous permettra de générer par rapport à l'équation (3.1), pour chaque métal donc chaque  $S$  un vecteur de  $U$  dont la dimension correspond à notre intervalle de prédiction, tout en gardant une relation de corrélation entre les vecteurs  $U$  pour prendre en compte l'interrelation entre les différents prix des métaux.

Nous devons également spécifier la taille de la simulation (500 itérations par exemple) et donc la dimension des vecteurs de nombres aléatoires. Ceci nous permettra ensuite de trier les résultats des simulations par ordre descendant et d'en tirer des quantiles qui seront nos scénarios. On aura ainsi par exemple (figure 3.3 ci-dessus) un scénario (A) correspondant à un quantile 80% de notre simulation et un scénario (C) correspondant à un quantile 10%.

**Étape 2 :** Le modèle codé sous le langage de programmation VBA d'Excel prend les différentes données entrées à l'étape 1 d'initialisation et calculera successivement :

- Des vecteurs de nombres aléatoires qui après application de la factorisation de Cholesky deviendront des vecteurs  $U$  de nombres aléatoires multivariés.
- Puis en appliquant notre équation (3.1), on calculera les différents  $S_t$  (prix futurs) selon plusieurs scénarios grâce aux vecteurs aléatoires  $U$  utilisés dans la simulation de Monte Carlo. Les simulations successives, c'est-à-dire l'application des différents  $U$  dans notre équation (3.1), nous permettront d'obtenir un résultat final résumé par le graphique ci-dessous :

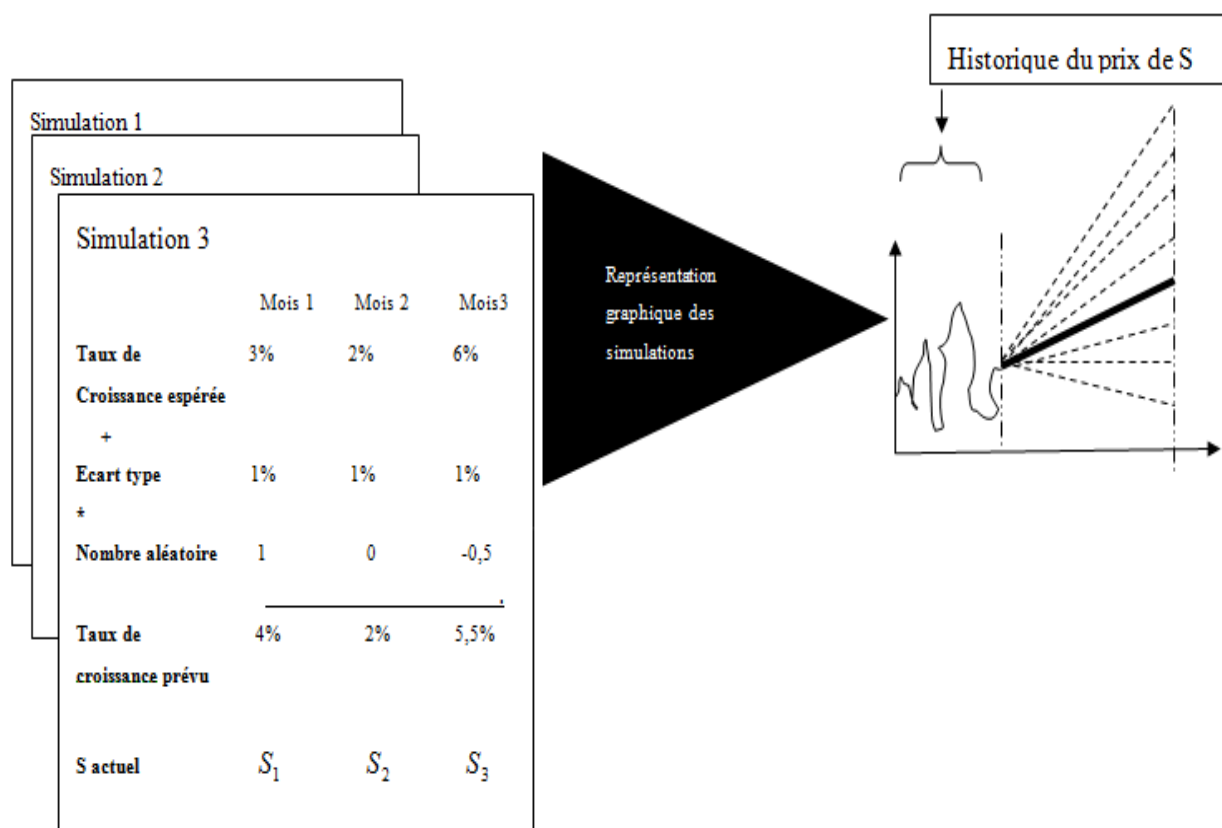


Figure 3-6 Schématisation des résultats de la simulation pour un actif

La différence entre les scénarios dans la figure ci-dessus provient des nombres aléatoires qui perturbent le taux de croissance prévu.

**Étape 3 :** Cette étape est celle d'une vérification manuelle des données obtenues pour s'assurer de leur cohérence en l'occurrence par exemple que pour tous les mois les valeurs prédites selon les scénarios correspondants aux quantiles 80% ont bien des valeurs supérieures à celles des quantiles inférieurs.

**Étape 4 :** Cette dernière étape cherche à reconstruire en fonction des prévisions précédentes des prix des matières premières, une prévision de l'indice synthétique censé représenter le WPU10. On raisonne donc d'une manière à rebours selon le graphique ci-dessous. Cependant comme chaque sous-composant du WPU10 a un poids spécifique dans l'indice, la recomposition de l'indice doit se faire selon ce poids.

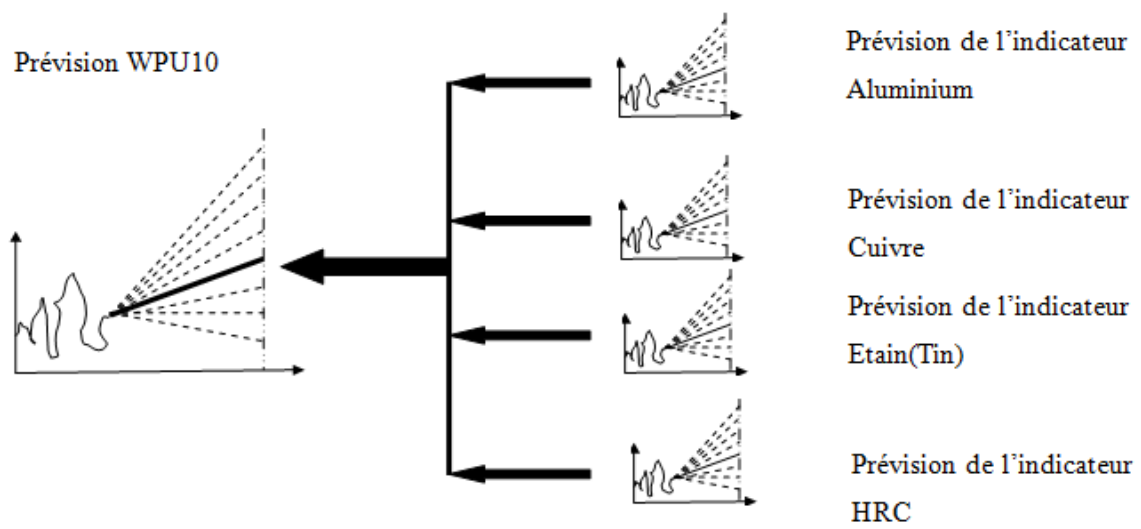


Figure 3-7 Recomposition du WPU10 selon le poids de chaque métal dans l'indice



## CHAPITRE 4 ANALYSE DES DONNÉES ET PRÉSENTATION DES RÉSULTATS

### 4.1 Décomposition de la variance du WPU10.

Le cadre de notre étude s'effectue sur la période allant de décembre 2003 à juin 2010, ce qui nous permettra de couvrir l'indice sous quatre types de régimes du marché :

- déc. 2003-déc. 2004 : marché en croissance lente
- janvier 2005-juillet 2008 : hyperinflation du marché
- août 2008-novembre 2009 : récession
- novembre 2009-juin 2010 : reprise lente du marché

On a introduit avec la formule (2.4) le principe d'indice de sensibilité connue sous la dénomination de « Standardized Regression Coefficient » ou SRC, pour pouvoir expliquer la contribution dans la variance du WPU10 de ses sous-composants.

Ainsi en posant :  $\beta_i$  poids du sous-composant i

$\sigma_i$  l'écart type de la croissance mensuelle du prix du sous-composant i

On cherchera à établir la contribution en termes d'effet de premier ordre de chaque sous-composant dans la variance de l'indice principal. Soit donc un  $SRC_i$  contribution du sous-composant i dans la variance de l'indice mère tel que :

$$SRC_i = \frac{\beta_i^2 * \sigma_i^2}{\sum_i \beta_i^2 * \sigma_i^2}$$

La décomposition de la variance du WPU10, tirée des informations sur sa composition fournies par le BLS ainsi que les données historiques de ses sous composants, nous permet de classer les sous-composants du WPU10 selon 2 catégories regroupés dans les 2 tableaux ci-dessous :

Tableau 4-1 Sous-composants du WPU10 les plus influents dans la contribution à sa variance.

Rang dans l'échelle de classification de parenté	Classification par code des PPIs (WPU)	Description	Poids dans WPU10	Ecart type historique du prix (sigma)	Poids*2*Sigma^2	SRC
0	10	Metals & Metal Products	100			
1	101	Iron and steel	26,521			
4	101211	Carbon steel scrap	3,109	14,51%	0,0020357%	43,88%
6	10170301	Hot rolled sheet and strip, incl. tin mill	3,68	6,98%	0,0006601%	14,23%
6	10170401	Hot rolled bars, plates, and structural shape	3,771	4,03%	0,0002312%	4,98%
6	10170601	Steel pipe and tube	3,72	3,10%	0,0001329%	2,86%
6	10170701	Cold rolled sheets and strip	3,2	4,49%	0,0002065%	4,45%
1	102	Nonferrous metals	22,567			
2	1021	Nonferrous metal ores	2,834	6,19%	0,000308%	6,63%
4	102201	Primary nonferrous metals	3,037	5,70%	0,000300%	6,46%
2	1023	Nonferrous scrap	2,059	6,41%	0,000174%	3,75%
4	102502	Copper and brass mill shapes	1,906	5,76%	0,000121%	2,60%
2	1026	Nonferrous wire and cable	3,965	3,44%	0,000186%	4,01%
2	1074	Struct., arch., pre-eng. metal products	8,062	1,28%	0,000106%	2,30%
<b>Total</b>			<b>39,343</b>	<b>1,95%</b>		<b>96,16%</b>

Tableau 4-2 Sous-composants du WPU10 les moins influents sur sa variance.

Rang dans l'échelle de classification de parenté	Classification par code des PPIs (WPU)	Description	Poids dans WPU10( %)	Ecart type historique du prix (sigma)	Poids^2*Sigma^2	SRC
0	10	Metals & Metal Products	100			
1	101	Iron and steel	26,521			
2	1015	Foundry and forge shop products	4,434	1,22%	0,0000294%	0,63%
1	102	Nonferrous metals	22,567			
4	102501	Aluminum mill shapes	3,251	2,07%	0,000045%	0,98%
2	1028	Nonferrous foundry shop products	2,732	0,90%	0,000006%	0,13%
2	1031	Metal cans and can components	2,487	1,24%	0,000010%	0,21%
1	104	Hardware	3,465	0,44%	0,000002%	0,05%
2	1071	Metal doors, sash, and trim	2,304	0,60%	0,000002%	0,04%
2	1073	Sheet metal products	3,384	0,93%	0,000010%	0,21%
1	108	Miscellaneous metal products	19,254	0,43%	0,000069%	1,48%
1	109	Metal treatment services	3,904	0,58%	0,000005%	0,11%
<b>Total</b>			<b>45,215</b>	<b>0,34%</b>		<b>3,84%</b>

Ces deux classifications nous amènent à la même conclusion, à savoir qu'une compréhension de la variance du WPU10 et une capacité à le prédire passe nécessairement par une meilleure compréhension de ses sous-composants les plus volatils. C'est pourquoi nous avons cherché à

comprendre ces indices en les associant à des actifs financiers dont nous avons une meilleure compréhension d'une part et pour lesquels nous disposons en temps réel et d'une manière continue de l'information d'autre part. En clair pour comprendre WPU10 au moyen de sa variance on cherchera à construire un indice synthétique fait d'autres éléments que l'on connaît mieux mais qui ont un comportement similaire à ceux du WPU10 et de ses sous-composants.

## **4.2 Identification des indicateurs avancés des sous composants du WPU10**

Afin de pouvoir faire des prédictions du WPU10, la technique la plus directe qui s'offre à nous est une reconstruction de l'indice d'une manière synthétique en allant de son élément le plus simple au plus compliqué. Nous avons donc essayé de trouver des indicateurs répliquant le comportement des sous-composants du WPU10 notamment ceux avec le plus grand impact sur sa variance pour pouvoir effectuer leur prédiction et par la suite reconstruire en fonction de leurs poids dans WPU10 les valeurs prédites pour avoir une valeur espérée future du WPU10. Nous avons pour cela utilisé deux outils :

- Au niveau mathématique, pour étudier la relation entre l'indicateur et le sous-composant du WPU10 nous avons fait des études de corrélations et de régression pour voir si leurs relations sont déterministes. Ici on a un compromis entre un tamisage selon la méthode de Morris et une analyse de sensibilité globale.
- Au niveau du choix de l'indicateur, nous avons utilisé des indicateurs avancés qui sont des produits financiers liés au marché des métaux. Ce sont entre autres, les contrats futurs qui représentent les anticipations des investisseurs financiers quant à l'évolution dans un future proche ou lointain du prix d'un actif sous-jacent au contrat qui peut être une action, une option, une commodité, un taux de change, un taux d'intérêt, etc.

Dans ce qui suit nous montrons un exemple d'analyse faite sur un sous-composant du WPU10 soit le WPU101211. Puis, nous présentons un tableau récapitulatif présentant les résultats des analyses faites sur tous les sous-composants du WPU10. Cette manière de faire nous permettra d'alléger ce présent mémoire. Nous pensons qu'il est néanmoins utile de présenter notre démarche d'analyse sur un cas puisqu'ensuite pour les autres cas ce sera toujours la même démarche qui s'appliquera.

### 4.2.1 Un exemple d'analyse : le cas du WPU101211 ou Carbon steel scrap

Cet indice représente tous les matériaux recyclables fait d'acier au carbone. Nous avons essayé de le modéliser en régressant l'indice par rapport au prix de la tonne métrique de plusieurs métaux tel que : l'acier (HRC = Hot Rolled Coil), le cuivre, l'aluminium, l'étain, le zinc, etc. Il faut noter que lorsqu'on effectue ces régressions, on essaie également de voir s'il y a un décalage temporel dans la relation entre notre indice et le prix d'une tonne métrique de la matière première d'intérêt. Par exemple, on étudie si le WPU101211 a une meilleure corrélation au prix par tonne du cuivre décalé de 3 mois en avance ou en retard ou une meilleure relation avec une moyenne mobile 3 mois, 6 mois, etc. du prix par tonne du cuivre comparativement à une corrélation avec le prix par tonne du cuivre du mois. Nous vous présentons l'analyse qui nous a fourni les résultats les plus concluants dans le cas du WPU101211, soit donc une régression du WPU101211 par rapport à la moyenne mensuelle du prix de la tonne métrique du Hot Rolled Coil (HRC) ou acier :

Tableau 4-3 Analyse du WPU101211

Corrélation					
	WPU101211		Acier (Hot rolled Coil)		
WPU101211		1			
Acier (Hot rolled Coil)		0,845			1
Statistiques de la régression					
Multiple R					0,845
R Square					0,714
Adjusted R Square					0,710
Standard Error					63,164
Observations					79
Anova					
	df	SS	MS	F	Significance F
Regression	1	767094,473	767094,47	192,27	1,248E-22
Residual	77	307210,951	3989,75		
Total	78	1074305,42			
Résultats de la régression					
	Coefficients	Standard Error	t Stat	P-value	Lower 95%
Intercept	-32,53	29,1008	-1,118	0,267	-90,48
Acier (Hot rolled Coil)	0,59	0,0428	13,867	0,000	0,51

Du tableau ci-dessus, on constate que la corrélation entre les deux variables est forte et qu'une grande partie (71%) de la variance du WPU101211 est corrélée avec le prix de l'acier. Par

ailleurs, on peut établir une relation linéaire entre les 2 variables puisque le coefficient de la régression a une valeur p significative.

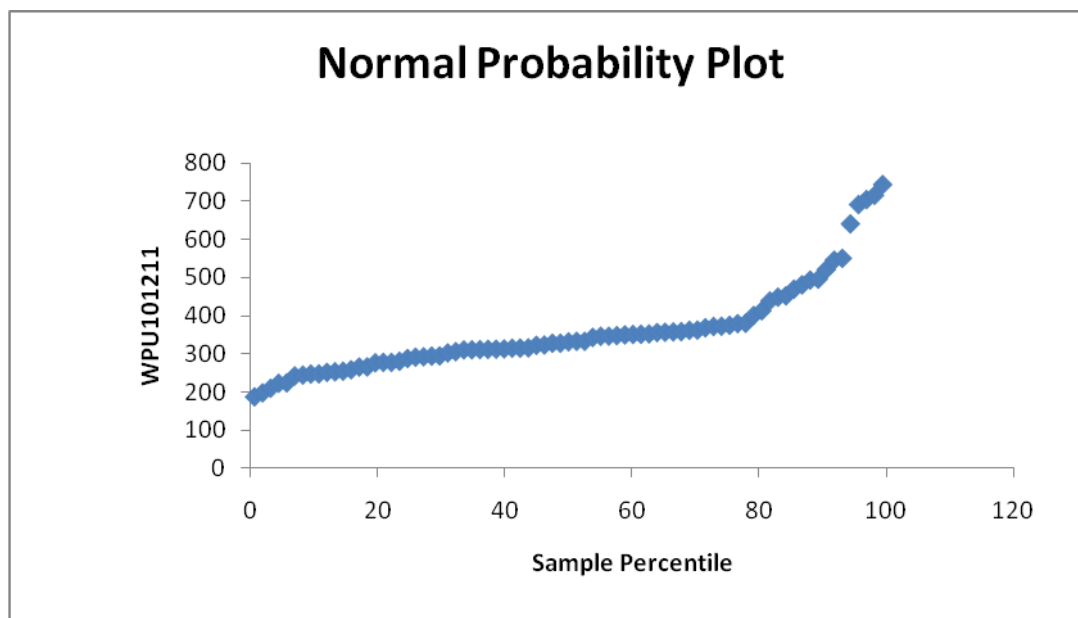


Figure 4-1 Distribution normale des résidus de la régression du WPU101211 sur le HRC

Cependant il est à noter que l'hypothèse de distribution normale des résidus de la régression n'est pas respectée.

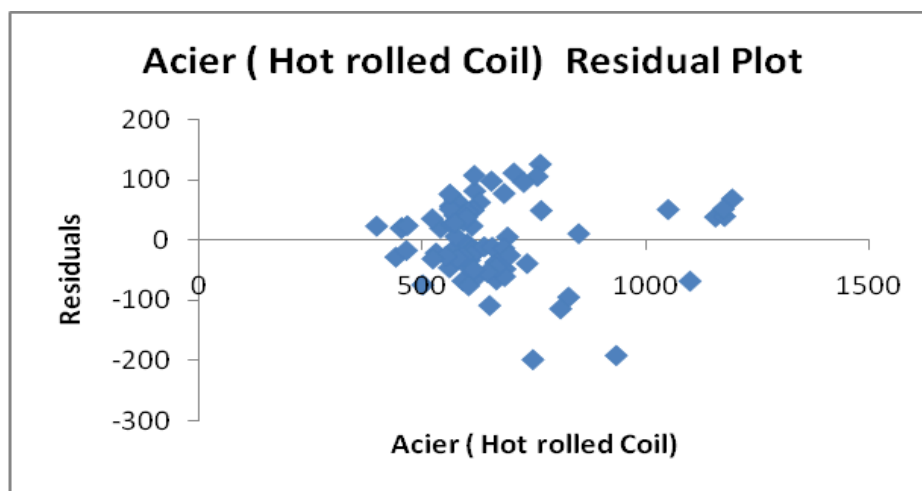


Figure 4-2 Répartition des résidus de la régression du WPU101211 sur le HRC

De plus les résidus de la régression bien qu'uniformément répartis, présentent parmi eux des valeurs extrêmes.

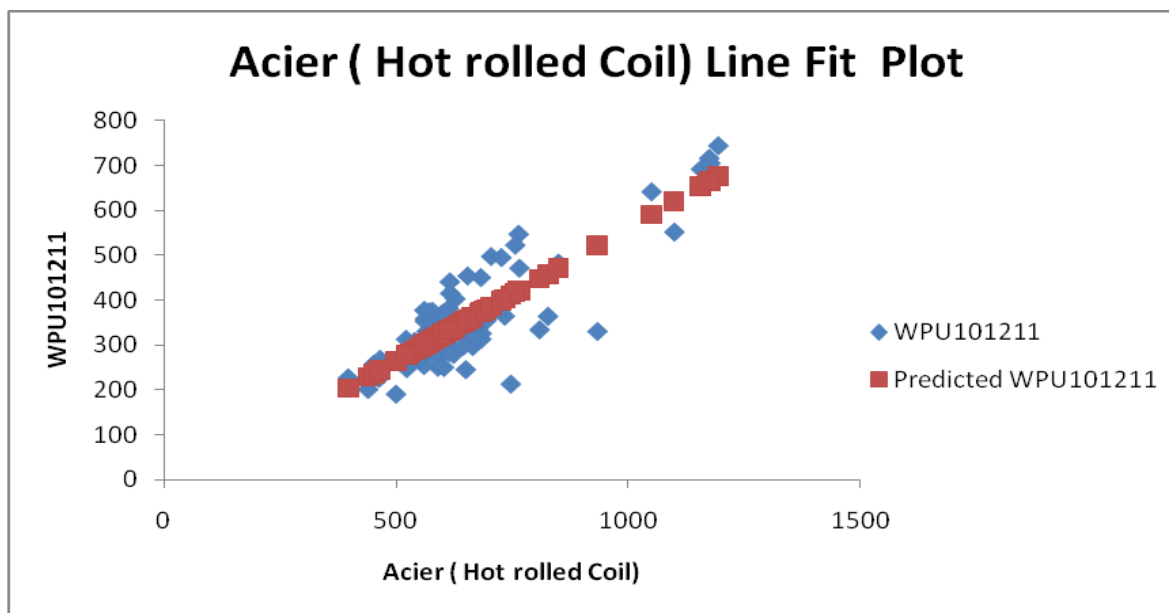


Figure 4-3 Comparaison des valeurs réelles et prédites du WPU101211

Malgré tout une comparaison des valeurs prédites par le modèle obtenu à partir de la régression avec les valeurs réelles nous porte à croire que le modèle est robuste malgré les défauts constatés.

#### 4.2.2 Tableau récapitulatif des analyses par sous-composant du WPU10

Dans le tableau 4.4 (ci-dessous) un récapitulatif des résultats des analyses telle que celle effectuée ci-dessus pour identifier des indicateurs au sous-composant WPU101211 est présenté, incluant les résultats précédents. On note que dans le tableau 4.4, il existe contrairement au cas précédent des cas d'analyses de régression multifactorielle parmi les résultats les plus concluants identifiés.

Tableau 4-4 Analyse des sous-composants de WPU10

Code commodité	Description (nom originel en anglais donné par le BLS)	Nom de l'indicateur	Corrélation avec indicateur	Coefficient directeur régression	Constante régression	R^2 Ajusté	Graph. Résidu normal?
10	Metals and metal products						
101211	Carbon steel scrap	HRC	0,84	0,59	0,0	0,71	Non
1015	Foundry and forge shop products	Aluminium Cuivre Moyenne Mobile-6mois (Cuivre MM-6) HRC	0,41 0,75 0,56	-0,02 0,01 0,03	146,4 146,4 146,4	0,75 0,75 0,75	Non (kurtosis)
10170301	Hot rolled sheet and strip, incl. tin mill	HRC	0,92	0,14	36,9	0,84	oui
10170401	Hot rolled bars, plates, and structural shape	HRC Cuivre Tin	0,80 0,68 0,86	0,07 0,00 0,00	78,4 78,4 78,4	0,82 0,82 0,82	Non (un peu de kurtosis)
10170601	Steel pipe and tube	Aluminium MM-6 HRC	0,61 0,67	0,02 0,08	70,4 70,4	0,54 0,54	Non
10170701	Cold rolled sheets and strip	étain Aluminium	0,74 0,79	0,00 0,03	59,7 0,0	0,74 0,00	Oui
1021	Nonferrous metal ores	Cuivre	0,96	0,03	51,7	0,93	
102201	Primary nonferrous metals	Cuivre	0,97	0,03	65,8	0,94	
1023	Nonferrous scrap	Régression exponentielle	0,00	0,00	0,0	0,00	
102501	Aluminum mill shapes	Aluminium Cuivre MM-6	0,91 0,94	0,01 0,00	114,2 114,2	0,94 0,94	Oui (Mais graph. des résidus avec les valeurs prédites pas satisfaisant)
102502	Copper and brass mill shapes	Cuivre	0,94	0,04	88,6	0,94	Non (Asymétrie à gauche + kurtosis)
1026	Nonferrous wire and cable	Cuivre	0,95	0,02	108,9	0,90	Non (Asymétrie à gauche + kurtosis)
1028	Nonferrous foundry shop products	Cuivre MM-6 Aluminium 6- MA	0,89 0,80	0,01 -0,02	156,8 0,0	0,82 0,82	Non (Kurtosis)
1031	Metal cans and can components	Aluminium Cuivre MM-6	0,24 0,61	-0,02 0,01	125,5 0,0	0,57 0,57	Oui
104	Hardware	Aluminium Cuivre MM-6	0,33 0,72	-0,02 0,01	175,0 0,0	0,71 0,71	Non
1071	Metal doors, sash, and trim	Aluminium Cuivre MM-6	0,39 0,75	-0,01 0,01	187,6 0,0	0,69 0,69	Non (Asymétrie & kurtosis)
1073	Sheet metal products	Aluminium Cuivre MM-6	0,43 0,78	-0,01 0,01	170,0 170,0	0,72 0,72	Oui (petite asymétrie à gauche)
1074	Structural, architectural, and preengineering metal products	Aluminium Cuivre MM-6 HRC	0,37 0,71 0,58	-0,03 0,01 0,04	163,0 163,0 163,0	0,72 0,72 0,72	Oui
108	Miscellaneous metal products	Aluminium Cuivre MM-6	0,33 0,72	-0,01 0,01	145,3 145,3	0,69 0,69	Oui
109	Metal treatment services	Aluminium	0,44	-0,01	144,5	0,79	Oui

Il ressort du tableau 4.4 que la plupart des sous-composants du WPU10 peuvent être indexés à des matières premières pour lesquelles il existe un marché futur. Ainsi, même si on ne connaît pas le comportement futur des sous-composants de WPU10, parce que leurs comportements sont similaires à ceux des matières premières telles que le cuivre, l'aluminium, etc. on peut les utiliser pour déduire leurs comportements futurs. Nous résumons une partie du tableau ci-haut, tel que chaque sous-composant du WPU10 est regroupé selon l'indicateur ou les indicateurs les plus influents.

Tableau 4-5 Les sous-composants de WPU10 regroupés par indicateur

Code commodité (WPU)	Pois dans WPU10	Description selon BLS	indicateur
1015	4,434	Foundry and forge shop products	Aluminium
10170701	3,2	Cold rolled sheets and strip	Aluminium
102501	3,251	Aluminum mill shapes	Aluminium
1031	2,487	Metal cans and can components	Aluminium
104	3,465	Hardware	Aluminium
1071	2,304	Metal doors, sash, and trim	Aluminium
1073	3,384	Sheet metal products	Aluminium
1074	8,062	Structural, architectural, and preengineering metal products	Aluminium
108	19,254	Miscellaneous metal products	Aluminium
109	3,904	Metal treatment services	Aluminium
10170601	3,72	Steel pipe and tube	Aluminium Moyenne Mobile-6mois
1028	2,732	Nonferrous foundry shop products	Aluminium Moyenne Mobile-6mois
10170401	3,771	Hot rolled bars, plates, and structural shape	Cuivre
1021	2,834	Nonferrous metal ores	Cuivre
102201	3,037	Primary nonferrous metals	Cuivre
102502	1,906	Copper and brass mill shapes	Cuivre
1026	3,965	Nonferrous wire and cable	Cuivre
1015	4,434	Foundry and forge shop products	Cuivre Moyenne Mobile-6mois
1028	2,732	Nonferrous foundry shop products	Cuivre Moyenne Mobile-6mois
1031	2,487	Metal cans and can components	Cuivre Moyenne Mobile-6mois
104	3,465	Hardware	Cuivre Moyenne Mobile-6mois
1071	2,304	Metal doors, sash, and trim	Cuivre Moyenne Mobile-6mois
1073	3,384	Sheet metal products	Cuivre Moyenne Mobile-6mois
1074	8,062	Structural, architectural, and preengineering metal products	Cuivre Moyenne Mobile-6mois
108	19,254	Miscellaneous metal products	Cuivre Moyenne Mobile-6mois
101211	3,904	Carbon steel scrap	Acier ou Hot Rolled Coil (HRC)
1015	3,72	Foundry and forge shop products	Acier ou Hot Rolled Coil (HRC)
10170301	2,732	Hot rolled sheet and strip, incl. tin mill	Acier ou Hot Rolled Coil (HRC)
10170401	3,771	Hot rolled bars, plates, and structural shape	Acier ou Hot Rolled Coil (HRC)
10170601	2,834	Steel pipe and tube	Acier ou Hot Rolled Coil (HRC)
1074	3,037	Structural, architectural, and preengineering metal products	Acier ou Hot Rolled Coil (HRC)
10170401	3,771	Hot rolled bars, plates, and structural shape	étain
10170701	3,2	Cold rolled sheets and strip	étain



Du tableau 4-5 ci-dessus on constate que tous les sous-composants influents du WPU10, dans leur ensemble, ont une variabilité qui est corrélée d'une manière plus ou moins forte au cuivre, à l'aluminium, à l'acier et à l'étain. Par ailleurs, on a établi à l'équation (3.4) qu'après avoir identifié une corrélation entre les sous-composants du WPU10 et les prix des matières premières, on chercherait à établir par transitivité une relation entre le WPU10 et ces mêmes matières premières. En l'occurrence pour chaque sous-composant du WPU10 on dispose au tableau 4-4 d'un modèle de régression le liant aux matières premières identifiés, et donc on peut recalculer les coefficients de ce modèle pour identifier une relation directe entre les matières premières du modèle et l'indice WPU10 sur la base du poids du sous-composant dans le WPU10. Ce qui, après avoir recalculer les coefficients de chaque modèle, nous laisserait à la fin avec un modèle de régression unique issu des modèles du tableau 4-4 liant le WPU10 et les quatre matières premières.

On a obtenu après calcul un nouveau modèle nous permettant de calculer le taux de variation mensuel de l'indice WPU10 sur la base de la variation des prix des quatre matières premières.

$$Y = 60,3\% + 8\%\Delta Cu. + 9,8\%\Delta Cu.-MM6 + 6,2\%\Delta Alu. + 0,7\%\Delta Alu.-MM6 + 2,1\%\Delta Etain + 0,13\Delta HRC \quad (4.1)$$

avec

Y : le taux de variation d'une période de WPU 10

$\Delta Cu$  : le taux de variation d'une période du prix d'une tonne de cuivre

$\Delta Cu-MM6$  : le taux de variation d'une période de la moyenne mobile 6 mois du prix du cuivre

$\Delta Alu$  : le taux de variation d'une période du prix d'une tonne d'aluminium

$\Delta Alu-MM6$  : le taux de variation d'une période de la moyenne mobile 6 mois du prix de l'aluminium

$\Delta Etain$  : le taux de variation d'une période du prix d'une tonne d'étain

$\Delta HRC$  : le taux de variation d'une période du prix d'une tonne d'acier

Sur la base de l'équation (4.1), calculer une valeur future d'une période du WPU10 devient facile puisqu'on n'aura qu'à multiplier notre valeur initiale de l'indice par le taux de variation Y sur la période. Par ailleurs comme les prix des matières premières sont disponibles sur plusieurs périodes futures on pourra calculer plusieurs taux de variations futurs du WPU10.

### 4.3 Test de la robustesse de l'indice synthétique : test historique

Afin de tester la robustesse de l'indice synthétique (indice WPU10 calculé sur la base de l'équation (4.1)), on a essayé de voir si notre indice synthétique répliquait réellement notre indice WPU10 en essayant de voir la qualité de la réplique sur une base historique. On a donc calculé :

- Premièrement, les taux de croissance mensuels de nos indicateurs de décembre 2003 à juin 2009.
- Puis, nous avons calculé les Y mensuels selon (4.1) de notre indice synthétique. Nous avons appliqué ces taux de croissances mensuel synthétiques obtenus à la valeur de l'indice WPU10 en commençant par la valeur de décembre 2003 pour obtenir ces valeurs synthétiques jusqu'à juin 2009.
- Enfin nous avons comparé les valeurs obtenues par modélisation sur l'intervalle décembre 2003 à juin 2009 aux valeurs réelles du WPU10 sur le même intervalle. Les résultats montrés dans le graphique ci-dessous sont plus que satisfaisants.

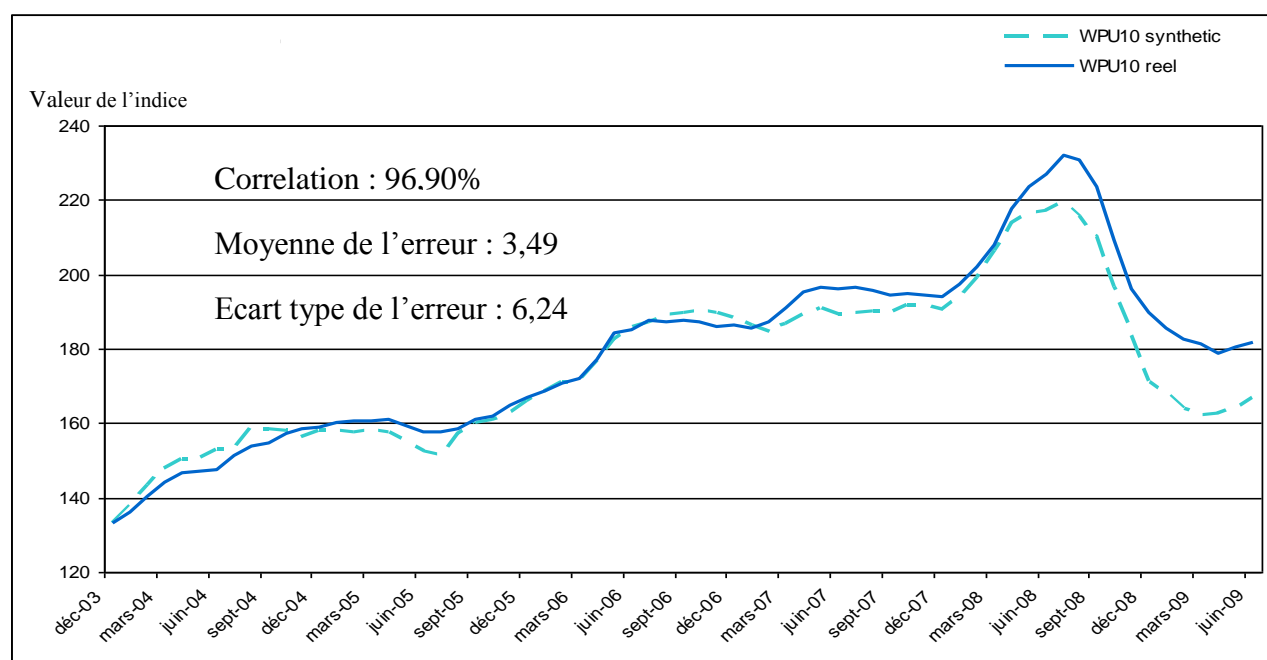


Figure 4-4 Test historique de la robustesse de l'indice synthétique

## 4.4 Un exemple détaillé de prévision : janvier 2010 à mars 2010

Nous allons vous montrer un exemple détaillé pas à pas de notre processus de prévision avec les sorties de notre programme et les explications nécessaires en accord avec ce que nous avons introduit auparavant pour une prévision de 3 mois et 5 itérations pour notre simulation de Monte Carlo

### 4.4.1 Initialisation : pourvoir notre modèle de données et instructions initiales

On énumère ci-dessous les étapes nécessaires pour initialiser notre processus de prévision.

- À chaque mois, on simule plusieurs valeurs de l'indice, soit dans notre cas ici 5 valeurs correspondant à nos 5 itérations. Par ailleurs, on indique à notre programme de simulation les quantiles qui nous intéressent lors de la simulation des différents prix pour pouvoir établir un intervalle de confiance autour de nos valeurs prédites. Ici dans cet exemple, on se concentre sur un intervalle de confiance délimité par nos quantiles 40% et 80%.
- On calcule notre matrice des corrélations sur la base des données historiques des prix des métaux soit : l'aluminium, le cuivre, l'étain et l'acier. On transforme par la suite cette dernière en utilisant la factorisation de Cholesky. Cette matrice est essentielle dans ce modèle de prévision puisqu'elle nous servira à générer nos nombres aléatoires multivariés :

Tableau 4-6 Matrice de corrélation des prix des matières premières

Commodités	Matrice de corrélation			
	Cuivre	Aluminium	Etain	Acier
Cuivre	1	0,44	0,33	0,2
Aluminium	0,44	1	0,45	0,56
Etain	0,33	0,45	1	0,34
Acier	0,2	0,56	0,34	1

Tableau 4-7 Matrice de corrélation après application de la décomposition de Cholesky

Commodités	Matrice de corrélation après décomposition de CHOLESKY			
	Cuivre	Aluminium	étain	Acier
Cuivre	1	0	0	0
Aluminium	0,44	0,897997773	0	0
Etain	0,33	0,339421777	0,880847806	0
Acier	0,2	0,525613776	0,108526396	0,8197269

- On donne au programme des données du prix des matières premières, ainsi que leurs variances (volatilités) et leurs taux de croissance.

Tableau 4-8 Prix commodités

Commodités	Unité	Date mis à jour	Source	Prix Actuel	Prix 3 Mois	Prix 15 Mois	Prix 27 mois
Cuivre	\$ USD/ton.	04-janv-10	LME	7464,00	7504,00	7555,00	7555,00
Aluminium	\$ USD/ton.	04-janv-10	LME	2247,00	2261,00	2388,00	2463,00
Etain	\$ USD/ton.	04-janv-10	LME	17450,00	17550,00	17585,00	17615,00
Acier	\$ USD/ton.	04-janv-10	NYMEX Steel Future	565,50	606,30	622,80	622,80

Tableau 4-9 Taux de croissance ( $\mu$ ) des prix des commodités

Commodités	Moyenne croissance mensuelle 3 mois	Moyenne croissance mensuelle 15 mois
Cuivre	0,178%	0,056%
Aluminium	0,207%	0,455%
Etain	0,190%	0,017%
Acier	2,322%	0,224%

Tableau 4-10 Variance commodités

Commodités	Volatilité actuelle	Volatilité 3 mois	Volatilité 15 mois
Cuivre	8,7%	8,7%	8,7%
Aluminium	11,5%	11,5%	11,5%
Etain	10,1%	11,0%	11,0%
Acier	13,0%	13,0%	13,0%

Ces variances correspondent aux variances historiques des taux de croissance mensuels des prix des différents métaux.

#### 4.4.2 Génération des nombres aléatoires

- Dans un premier temps on demande au programme de générer 3 matrices de nombres aléatoires distribués selon une loi normale (0,1), ces 3 matrices serviront à faire les simulations pour les trois mois de prévisions.

Matrice 1 (Mois 1)				
0,4146736	-1,10208475	-0,31073429	0,19373265	-0,38988598
1,7133893	1,02008086	-0,27356808	1,41380677	0,39079885
0,28258363	-1,5552517	-1,61287077	-0,82231855	0,9379756
0,06078853	-0,12008138	1,06844753	0,95191846	1,36729436

Matrice 2 (Mois 2)				
1,47959114	-1,18745359	0,32450551	-0,24304056	0,41090863
0,26088204	-0,30765217	0,26976744	0,03065679	-0,38033054
0,20007793	-0,20243787	0,30814413	-0,01244518	-0,18360145
1,59825663	0,87099633	-0,71907829	1,47958591	0,82106033

Matrice 3 (Mois 3)				
1,13717456	0,18786369	-2,07028502	0,55895686	-0,69131462
-1,05199707	1,72591705	0,93344981	-0,95847942	1,27665188
-0,80158633	0,76898563	0,55770341	-1,48595548	0,84526748
-0,05827632	0,85282827	-1,28537129	-0,54934358	-0,61594071

Chaque ligne de la matrice servira à générer un vecteur de nombres aléatoires multivariés pour une commodité et une seule. Par ailleurs, pour une ligne spécifique, chaque colonne de la matrice servira à générer un scénario, chaque cellule correspond donc à une valeur de  $U$  dans notre équation (3.1).

Des trois matrices ci-dessus le programme génère grâce à la matrice de Cholesky trois autres matrices de nombres aléatoires cette fois-ci multivariés. Vous remarquerez que les lignes des matrices ci-dessous sont classées par commodités contrairement à celles d'avant. La raison est qu'avant application de la matrice de Cholesky, n'importe quelle ligne des matrices aurait pu servir pour la simulation de n'importe quelle commodité indépendamment des autres commodités, ce qui n'est plus vrai après Cholesky.

Z Matrice de nombres aléatoire corrélés (Mois 1)					
Simulation	1	2	3	4	5
Cuivre	0,75525958	0,3206625	-1,67959756	-2,76547919	0,05592395
Aluminium	0,3558227	-0,14315315	-1,30132449	-0,67105799	0,00246579
étain	-0,63826685	1,46873888	-0,82868467	-0,15574074	-0,81139841
Acier	-0,95420488	-1,02182713	-1,00070437	0,72375408	-0,38703144

Z Matrice de nombres aléatoire corrélés (Mois 2)					
Simulation	1	2	3	4	5
Cuivre	1,47959114	-1,18745359	0,32450551	-0,24304056	0,41090863
Aluminium	0,88529159	-0,79875055	0,38503298	-0,07940812	-0,16073618
étain	0,75305233	-0,67460049	0,47007984	-0,08076011	-0,15521756
Acier	1,7648891	0,29281232	-0,34931146	1,17901125	0,53539439

Z Matrice de nombres aléatoire corrélés (Mois 3)					
Simulation	1	2	3	4	5
Cuivre	1,13717456	0,18786369	-2,07028502	0,55895686	-0,69131462
Aluminium	-0,44433422	1,63252969	-0,07268956	-0,61477137	0,84225211
étain	-0,68787867	1,32516815	0,12489096	-1,44977365	0,94974163
Acier	-0,46027319	1,72728002	-0,91655079	-1,00357571	0,11959356

#### 4.4.3 Statistiques des nombres aléatoires

Après avoir généré nos nombres aléatoires, le programme va générer des statistiques sur les nombres aléatoires multivariés. Ces statistiques serviront à avoir des sorties de nos simulations qui nous permettront d'identifier les intervalles de solutions qui nous intéressent conformément aux informations que l'on a fournies au programme initialement. Il peut paraître incongru de vouloir avoir des quantiles sur des vecteurs de matrices de petites dimensions comme c'est le cas dans notre exemple, mais il faut garder à l'esprit que dans l'application réelle de ce mémoire chaque matrice peut avoir un très grand nombre de colonnes où une telle information facilite la lecture des données. Les nombres ci-dessous nous donnent nos « U » de l'équation (3.1) qui nous intéressent par scénario.

Tableau 4-11 Statistiques nombres aléatoires multi variés

Statistiques des nombres aléatoires multi variés			
Moyenne	Mois 1	Mois 2	Mois 3
<b>Cuivre</b>	-23,89%	78,74%	-46,62%
<b>Aluminium</b>	66,08%	79,74%	7,02%
<b>Etain</b>	-27,73%	65,41%	60,99%
<b>Acier</b>	88,61%	25,28%	32,40%
Données quantile 0,4	Mois 1	Mois 2	Mois 3
<b>Cuivre</b>	-34,24%	81,95%	-64,65%
<b>Aluminium</b>	33,04%	52,72%	-9,30%
<b>Etain</b>	-66,33%	11,38%	43,93%
<b>Acier</b>	83,64%	12,04%	32,79%
Données quantile 0,8	Mois 1	Mois 2	Mois 3
<b>Cuivre</b>	23,79%	121,01%	1,73%
<b>Aluminium</b>	142,81%	129,03%	143,51%
<b>Etain</b>	85,76%	143,43%	95,08%
<b>Acier</b>	137,46%	71,92%	55,97%

#### 4.4.4 Calcul des taux de croissance mensuels selon les différents scénarios

Le tableau 4.12 ci-dessous, nous donne les taux de croissance sur des périodes de 3 mois et 15 mois de nos différentes commodités calculés sur la base de leurs prix affichés dans les différents marchés. Nous avons donc cherché, en utilisant l'équation (3.1), les taux affichés au tableau 4-9, les variances affichées au tableau 4-10 et les statistiques des nombres aléatoires du tableau 4-11 représentant nos différents scénarios (nos « U » de (3.1)), à calculer les taux de croissance mensuels des différentes commodités selon les différents scénarios (quantiles). Les résultats sont présentés ci-dessous tels que calculés par notre programme :

Tableau 4-12 Taux de croissance mensuels par commodité par scénario

Taux de croissance mensuel espéré	Mois 1	Mois 2	Mois 3
Cuivre	-1,890%	6,997%	-3,981%
Aluminium	7,837%	9,414%	1,266%
Etain	-2,851%	7,365%	6,708%
Acier	13,833%	5,606%	4,432%
Taux de croissance mensuel quantile 0,4	Mois 1	Mois 2	Mois 3
Cuivre	-2,787%	7,276%	-5,543%
Aluminium	4,022%	6,294%	-0,618%
Etain	-7,085%	1,439%	4,836%
Acier	13,187%	3,887%	4,484%
Taux de croissance mensuel quantile 0,8	Mois 1	Mois 2	Mois 3
Cuivre	2,239%	10,658%	0,207%
Aluminium	16,697%	15,106%	17,026%
Etain	9,598%	15,925%	10,446%
Acier	20,179%	11,664%	7,495%

#### 4.4.5 Prévision des différentes commodités

A cette étape on combine les données des tableaux 4-8 et 4-12. On applique les taux de croissance du tableau 4-12 aux prix affichés sur les marchés financiers du tableau 4-8. On obtient les prévisions des prix :



Tableau 4-13 Pr vision des commodit s

Pr�vision esp�r�e	Prix Actuel	Prix Mois 1	Prix Mois 2	Prix Mois 3
<b>Cuivre</b>	7464,00	7322,90	7835,32	7523,37
<b>Cuivre Moyenne Mobile 6 mois</b>	7464,00	7322,90	7579,11	7560,53
<b>Aluminium</b>	2247,00	2423,11	2651,22	2684,80
<b>Aluminium Moyenne Mobile 6 mois</b>	2247,00	2423,11	2537,16	2586,38
<b>Etain</b>	17450,00	16952,43	18201,02	19421,86
<b>Acier</b>	565,50	643,72	679,81	709,94
Pr�vision quantile 0,4	Prix Actuel	Prix Mois 1	Prix Mois 2	Prix Mois 3
<b>Cuivre</b>	7464,00	7255,97	7783,89	7352,45
<b>Cuivre Moyenne Mobile 6 mois</b>	7464,00	7255,97	7519,93	7464,10
<b>Aluminium</b>	2247,00	2337,38	2484,50	2469,14
<b>Aluminium Moyenne Mobile 6 mois</b>	2247,00	2337,38	2410,94	2430,34
<b>Etain</b>	17450,00	16213,60	16446,95	17242,33
<b>Acier</b>	565,50	640,07	664,95	694,76
Pr�vision quantile 0,8	Prix Actuel	Prix Mois 1	Prix Mois 2	Prix Mois 3
<b>Cuivre</b>	7464,00	7631,09	8444,38	8461,83
<b>Cuivre Moyenne Mobile 6 mois</b>	7464,00	7631,09	8037,74	8179,10
<b>Aluminium</b>	2247,00	2622,18	3018,29	3532,19
<b>Aluminium Moyenne Mobile 6 mois</b>	2247,00	2622,18	2820,24	3057,56
<b>Etain</b>	17450,00	19124,90	22170,46	24486,40
<b>Acier</b>	565,50	679,61	758,88	815,76

#### 4.4.6 Pr vision du WPU10

On sait sur la base de (4.1) que 60,3% du taux de croissance du WPU10 est constant. On a donc pond r  pour chaque mois chaque multiplicateur de croissance (chiffre obtenu en additionnant 1 au taux de croissance mensuel) de chaque commodit  par son poids dans WPU10 selon (4.1), puis nous avons calcul  leur somme en y additionnant la constante de 60,3%. On a donc obtenu le tableau donnant les multiplicateurs de croissance mensuels du WPU10 :

Tableau 4-14 Multiplicateur de croissance du WPU10

Prévision espérée	Multiplicateur de croissance mois 1	Multiplicateur de croissance mois 2	Multiplicateur de croissance mois 3
Cuivre	7,84%	8,56%	7,68%
Cuivre Moyenne Mobile 6 mois	9,57%	10,09%	9,73%
Aluminium	6,69%	6,78%	6,28%
Aluminium Moyenne Mobile 6 mois	0,75%	0,73%	0,71%
Etain	1,99%	2,20%	2,19%
Acier	14,76%	13,69%	13,54%
Constante	<b>60,3%</b>	<b>60,3%</b>	<b>60,3%</b>
WPU10	1,019	1,024	1,005

Prévision quantile 0,4	Multiplicateur de croissance mois 1	Multiplicateur de croissance mois 2	Multiplicateur de croissance mois 3
Cuivre	7,77%	8,58%	7,55%
Cuivre Moyenne Mobile 6 mois	9,48%	10,11%	9,68%
Aluminium	6,45%	6,59%	6,16%
Aluminium Moyenne Mobile 6 mois	0,72%	0,72%	0,70%
Etain	1,91%	2,08%	2,15%
Acier	14,68%	13,47%	13,55%
Constante	60,3%	60,3%	60,3%
WPU10	1,013	1,019	1,001

Prévision quantile 0,8	Multiplicateur de croissance mois 1	Multiplicateur de croissance mois 2	Multiplicateur de croissance mois 3
Cuivre	8,18%	8,85%	8,01%
Cuivre Moyenne Mobile 6 mois	9,97%	10,27%	9,92%
Aluminium	7,23%	7,14%	7,26%
Aluminium Moyenne Mobile 6 mois	0,81%	0,75%	0,75%
Etain	2,25%	2,38%	2,27%
Acier	15,58%	14,48%	13,94%
Constante	60,3%	60,3%	60,3%
WPU10	1,044	1,042	1,025

En appliquant donc ce nouveau multiplicateur à WPU10, on obtient les prévisions :

Tableau 4-15 Prévision de l'indice WPU10

Prévision de l'indice WPU10	Nov.-09	Déc.-09	Janv.-10	Févr.-10	Mars-10
Prévision espérée	193	196	193,212	192,133	190,242
Prévision quantile 40%	193	196	187,098	178,291	170,221
Prévision quantile 80%	193	196	188,908	183,211	177,098
WPU10 valeur réelle réalisée	193	196	191,892	190,412	187,132
Différence espérée/réelle			196,458	200,340	202,514
Différence quantile 80% /réelle			200,5	200,5	204,5

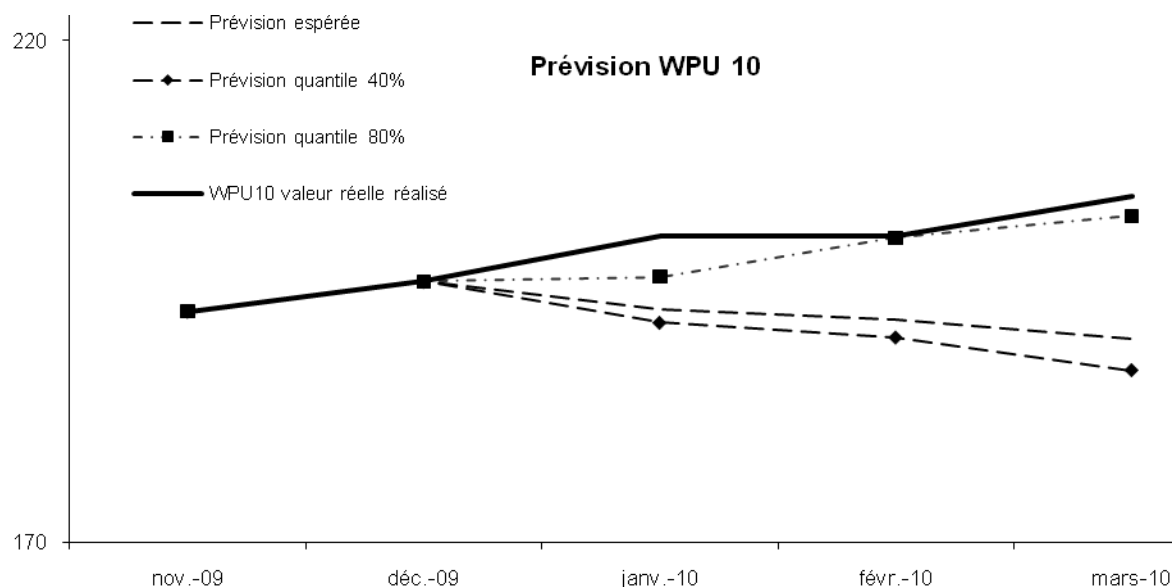


Figure 4-5 Prédiction de l'indice WPU10

On remarque que les valeurs du scénario quantile 80% sont des bonnes prévisions de notre valeur réelle de WPU10 qui s'est réalisée. On notera également que l'utilisation des intervalles de prévision est une bonne méthode puisque la valeur du quantile le plus bas (40%) nous indique que dans les multiples scénarios du modèle on peut voir par exemple qu'au pire des cas notre indice aurait perdu en janvier 2010 par rapport à sa valeur de décembre 2009 4,54% donc dans une économie baissière l'utilisateur du modèle peut se demander par exemple de combien en moyenne le prix des métaux peuvent bouger dans un marché baissier au pire des cas et donc le 4,54% peut être une information utile.

## 4.5 Les résultats de la prévision longue janvier 2010 à juin 2010

Nous avons refait la simulation, cette fois-ci pour une prévision de 6 mois avec 50 itérations au lieu de 5. Les détails sont tel que présentés dans l'annexe 5. On présente ci-dessous la prévision obtenue.

### 4.5.1 Prédiction du WPU10 selon 5 scénarios

On présente ci-dessous les prévisions de notre indice d'intérêt, à savoir le WPU10, selon 5 scénarios qui correspondent aux quantiles 5%, 16%, 40% et 80% ainsi que la moyenne de nos simulations pour chaque période. On notera que la valeur du WPU10 réelle et celle de la valeur

espérée sous forme de moyenne sont très proches et que le fait d'avoir WPU10 dans des scénarios peut nous donner un intervalle de confiance et une idée quant aux évolutions futures probables de l'indice.

Tableau 4-16 Prévision du WPU10

Prévision WPU10	nov.-09	déc.-09	janv.-10	févr.-10	mars-10	avr.-10	mai-10	juin-10
Prévision espérée	193	196	197,197	195,882	197,286	197,962	197,739	198,786
Prévision quantile 0,05	193	196	185,375	174,718	166,846	157,598	150,195	144,466
Prévision quantile 0,16	193	196	190,315	183,193	177,717	171,095	165,403	160,915
Prévision quantile 0,4	193	196	196,347	194,180	193,902	193,128	190,326	188,826
Prévision quantile 0,6	193	196	203,710	208,511	216,213	224,227	230,405	237,209
WPU10 valeur réelle réalisée	193	196	200,5	200,5	204,5	210,5	210,4	206,5
Différence réelle espérée			-3,303	-4,618	-7,214	-12,538	-12,661	-7,714

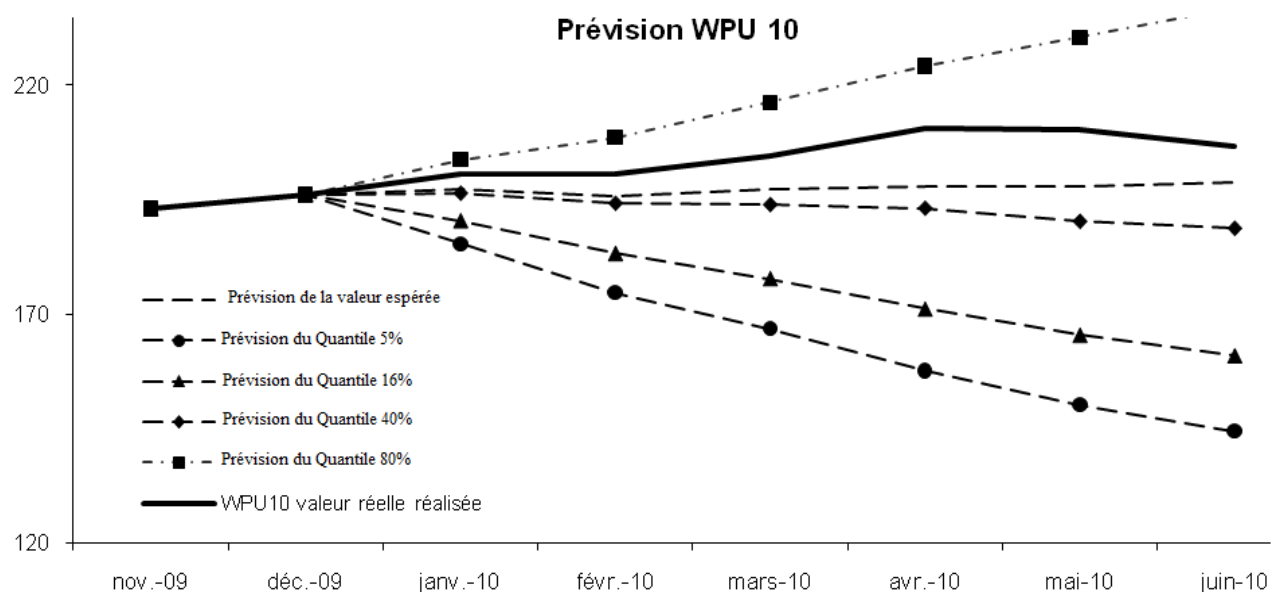


Figure 4-6 Prévision du WPU10

## 4.6 Conclusion

Ainsi on a établi que sur la base d'un test historique que le modèle de prédiction est un bon indicateur de l'évolution future du WPU10 avec une corrélation de 96,9%, une erreur moyenne de 3,49 représentant 1,69% de la valeur réelle du WPU10 en juin 2010 et ayant un écart type de 6,24. Par ailleurs pour un test effectué en prévision au mois de janvier 2010 pour une période de 6 mois on obtient une erreur moyenne de 8,008 soit 3,87% de la valeur de juin du WPU10. Notre modèle de prévision semble donc robuste. L'exercice d'analyse de sensibilité qui nous a permis de construire notre modèle dans le cas des marchés financiers nous semble être une bonne option. Cependant, dans la mesure où notre indice est une série temporelle on peut se poser la question de savoir si l'application d'autres méthodes d'analyse ne serait pas aussi efficace ou plus rapide? Dans le chapitre suivant, on développe les méthodes d'analyse de séries chronologiques pour répondre à cette question.

## CHAPITRE 5 MODÈLE D'ANALYSE DE SÉRIES TEMPORELLES

L'indice WPU10 ainsi que ses sous-composants sont également des séries temporelles, ce sont des suites de valeurs numériques représentant l'évolution de quantités spécifiques au cours du temps. À cet égard, il existe dans le domaine scientifique [13] des modèles consacrés exclusivement à l'étude de ce type de données. Nous montrons leur application dans le cas du WPU10 puis nous comparons un de ces modèles au modèle de diffusion.

### 5.1 Rappel sur la théorie

#### 5.1.1 Le modèle

On modélise le comportement d'une série temporelle au moyen d'un modèle autorégressif, on cherche à établir pour une variable  $y_t$  (variable observée au temps  $t$ ) une relation du type :

$$y_t = \phi y_{t-1} \quad (5.1)$$

où  $t = 1, 2, \dots, T$  et  $y_0$  est l'observation de départ.

La valeur de  $y_t$  peut s'écrire comme :

$$y_t = \phi^t y_0 \quad (5.2)$$

Cette dernière formulation permet de visualiser rapidement l'évolution de  $y_t$  sous différentes conditions. Plus spécifiquement, la valeur de  $\phi$  devient centrale quant à l'évolution de la récurrence : quand  $|\phi| < 1$ , les valeurs successives de  $y$  tendent vers 0 qui correspond à une valeur d'équilibre de long terme ou une valeur stable (qui se répète). Quand  $\phi$  est négatif, les valeurs de  $y$  oscillent. Si  $\phi > 1$ , les valeurs de  $y$  explosent. On pourra noter que la variable  $y_t$  peut avoir une part constante dans son comportement; dans ce cas (5.1) s'écrit :

$$y_t = \delta + \phi y_{t-1} \quad (5.3)$$

Il y a également le cas où la variable  $y_t$  a une partie stochastique et le comportement pourra être modélisé sous la forme :

$$y_t = \delta + \phi y_{t-1} + \varepsilon_t \quad (5.4)$$

où  $\varepsilon_t$  est généralement admis comme un bruit blanc qui suit souvent une distribution normale  $N(0, \sigma^2)$ . On montre [13] que si  $|\phi| < 1$ , alors :

**Moyenne :** 
$$\mu = \frac{\delta}{1 - \phi} \quad (5.5)$$

**Variance :** 
$$\text{var} = \frac{\sigma^2}{1 - \phi^2} \quad (5.6)$$

### 5.1.2 Importance de la stationnarité dans ce type de modélisation

Dans le cadre de l'étude de séries temporelles, un élément primordial dans l'analyse de leur comportement est le fait de savoir si elles suivent un processus stationnaire. On entend par là le fait que la structure du processus sous-jacent supposé évolue ou non avec le temps. Si la structure reste la même, le processus est dit alors stationnaire. Il est important de savoir cela puisque si les moments de la variable évoluent avec le temps les modèles autorégressifs deviennent difficilement applicables voire même impossible à appliquer. Il existe plusieurs critères tels que la tendance, la saisonnalité ou une corrélation des résidus de la différence des valeurs réelles et théoriques de la variable qui font qu'il y a ou pas de stationnarité. Une manière d'y remédier est de prendre la différence des valeurs originales de la variable pour enlever la tendance, par exemple.

## 5.2 Calibrage du modèle autorégressif au cas du WPU10

### 5.2.1 Analyse initiale

Nous étudions l'indice WPU10 sur la période décembre 2003 à décembre 2009.

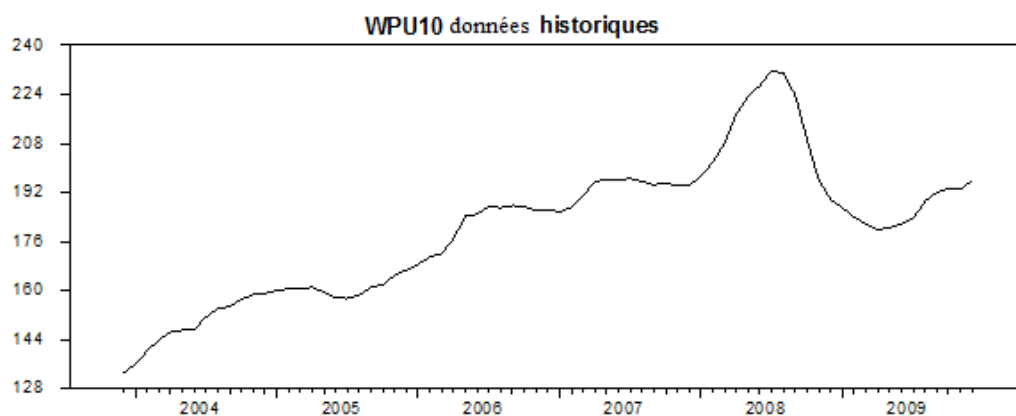


Figure 5-1 Données historiques du WPU10

On peut visuellement constater qu'il semblerait y avoir une tendance dans le WPU10. Ce que le calcul mathématique nous confirme :

Linear Regression - Estimation by Least Squares

Dependent Variable WPU10

Monthly Data from 2003:12 to 2009:12

Usable Observations	73	Degrees of Freedom	71
---------------------	----	--------------------	----

Centered R**2	0.645205	R Bar **2	0.640208
---------------	----------	-----------	----------

Uncentered R**2	0.994452	T x R**2	72.595
-----------------	----------	----------	--------

Mean of Dependent Variable	180.23698630
----------------------------	--------------

Std Error of Dependent Variable	22.87471619
---------------------------------	-------------

Standard Error of Estimate	13.72086024
----------------------------	-------------

Sum of Squared Residuals	13366.602409
--------------------------	--------------

Regression F (1, 71)	129.1157
----------------------	----------

Significance Level of F	0.00000000
-------------------------	------------

Durbin-Watson Statistic	0.073473
-------------------------	----------



Variable	Coeff.	Std Error	T-Stat	Signif
*****				
1. Constant	138.66897445	3.99518404	34.70903	0.00000000
2. TREND	0.86600025	0.07621289	11.36291	0.00000000

On peut constater qu'une régression par rapport à une variable tendance temps de WPU10 et une constante nous donne des coefficients significatifs et un  $R^2$  plutôt acceptable. Le même résultat peut être constaté visuellement :

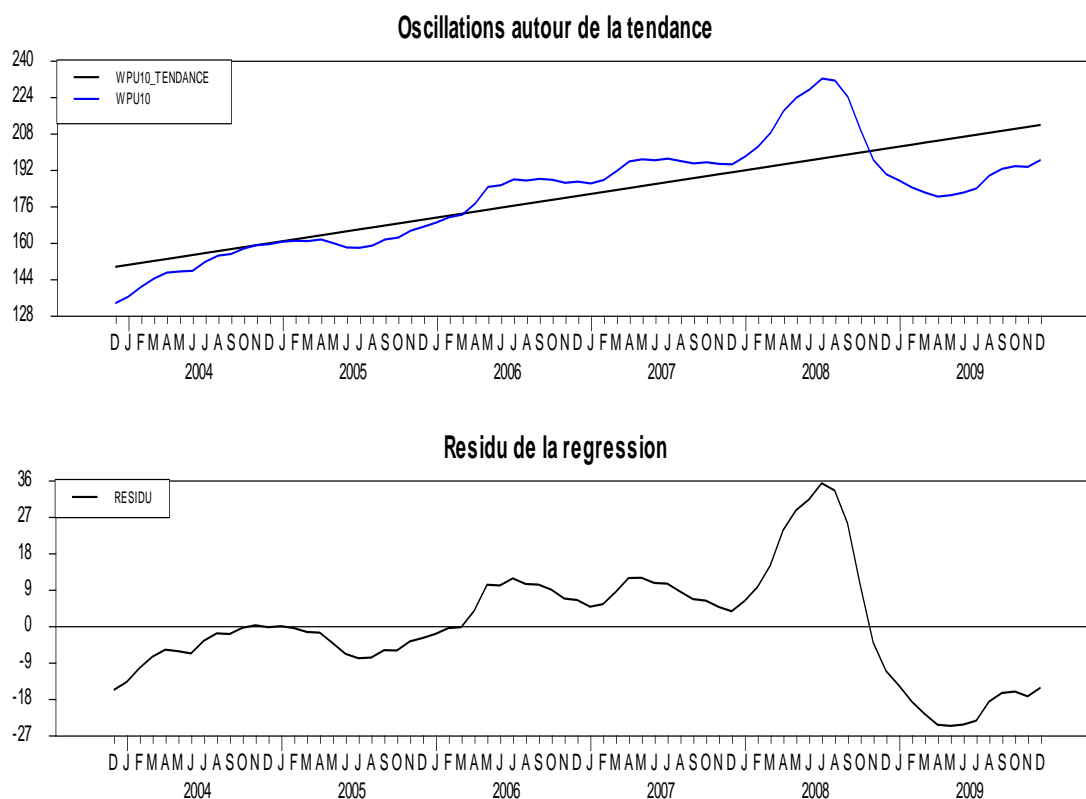


Figure 5-2 Oscillation du WPU10 par rapport à sa tendance

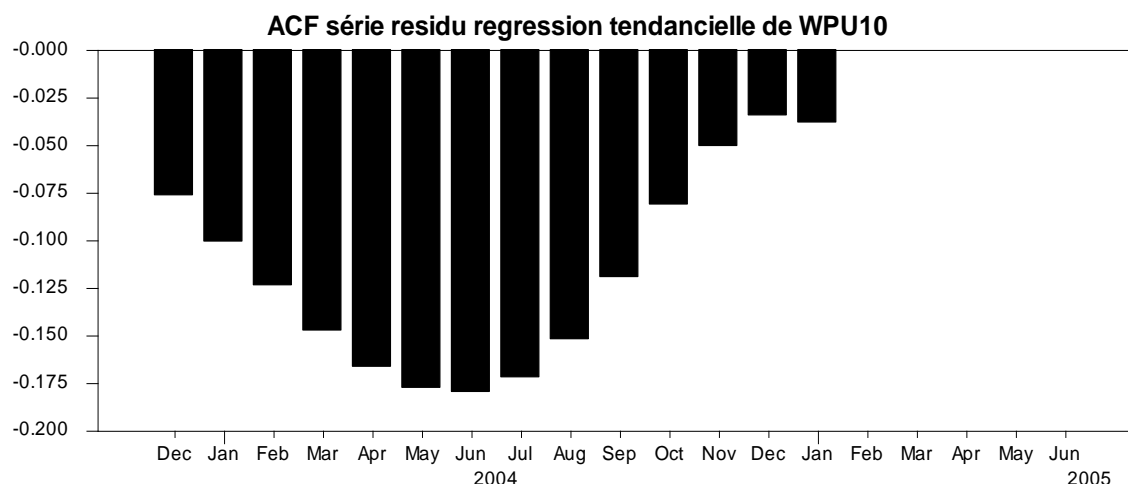


Figure 5-3 Autocorrélations des résidus de la régression de WPU10 par rapport à sa tendance

#### Ljung-Box Q-Statistics

Q (4-0) = 198.8007. Significance Level 0.00000000

Q (8-0) = 224.8784. Significance Level 0.00000000

Q (12-0) = 226.4948. Significance Level 0.00000000

Q (16-0) = 235.6284. Significance Level 0.00000000

Q (20-0) = 245.6299. Significance Level 0.00000000

Q (24-0) = 246.9211. Significance Level 0.00000000

Il apparaît que notre série n'est probablement pas stationnaire, puisque d'une part la dispersion des résidus de la régression est trop élevée et que d'autre part les auto corrélations de résidus de la régression tendancielle bien que négatives et faibles mettent du temps à s'estomper. On notera également que le test de Ljung-Box qui teste l'hypothèse de la nullité de la covariance des résidus rejette «  $H_0 : \text{Cov} = 0$  ». Nous allons essayer de voir s'il n'y a pas de saisonnalité en créant des variables de saisonnalité pour voir si une régression par rapport à ces dernières est significative. Les résultats de la régression sont tels que présentés ci-dessous :

## Linear Regression - Estimation by Least Squares

Dependent Variable WPU10

Monthly Data from 2003:12 to 2009:12

Usable Observations	73	Degrees of Freedom	60
Centered R**2	0.662140	R Bar **2	0.594568
Uncentered R**2	0.994716	T x R**2	72.614
Mean of Dependent Variable	180.23698630		
Std Error of Dependent Variable	22.87471619		
Standard Error of Estimate	14.56514253		
Sum of Squared Residuals	12728.602608		
Regression F (12, 60)	9.7990		
Significance Level of F	0.00000000		
Durbin-Watson Statistic	0.066689		

Variable	Coeff	Std Error	T-Stat	Signif
*****				
1. TREND	0.85914588	0.08173898	10.51085	0.00000000
2. T1	135.52339380	6.90730626	19.62030	0.00000000
3. T2	136.28091459	6.94925531	19.61087	0.00000000
4. T3	137.52176871	6.99190829	19.66870	0.00000000
5. T4	140.04595616	7.03525239	19.90632	0.00000000
6. T5	141.42014361	7.07927492	19.97664	0.00000000
7. T6	141.14433107	7.12396331	19.81261	0.00000000
8. T7	142.58518519	7.16930509	19.88829	0.00000000
9. T8	142.85937264	7.21528795	19.79954	0.00000000
10. T9	141.81689342	7.26189972	19.52890	0.00000000
11. T10	139.27441421	7.30912835	19.05486	0.00000000
12. T11	136.59860166	7.35696198	18.56726	0.00000000
13. T12	133.77528345	6.76016423	19.78876	0.00000000

On constate que nos variables saisonnalités T1, T2, etc. sont significatives.

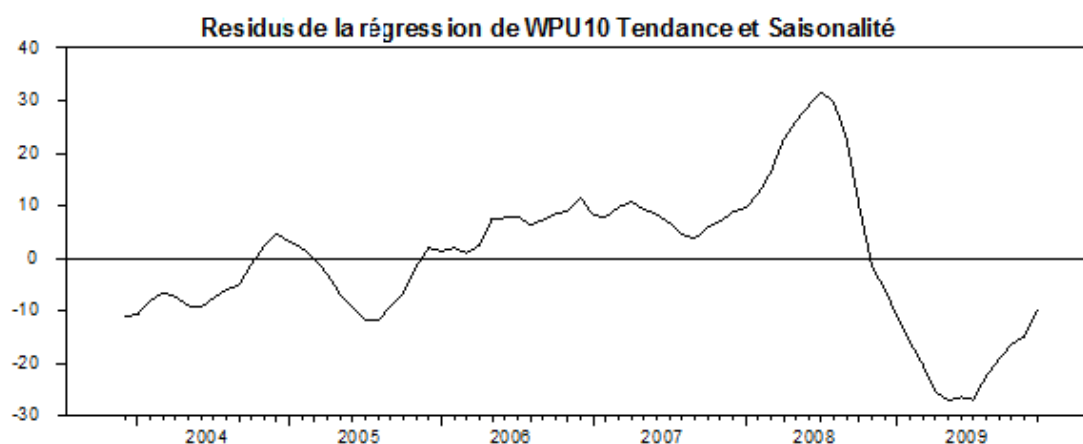


Figure 5-4 Résidus de la régression de WPU10 Tendance et saisonnalité

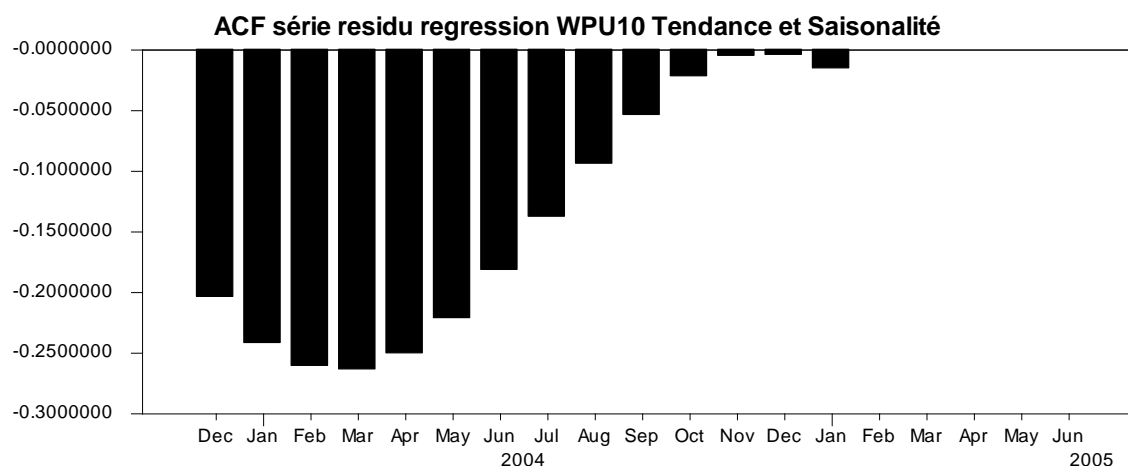


Figure 5-5 Autocorrélations des résidus de la régression de WPU10 Tendance et saisonnalité

#### Ljung-Box Q-Statistics

Q (4-0) = 204.7492. Significance Level 0.00000000

Q (8-0) = 234.5721. Significance Level 0.00000000

Q (12-0) = 245.5015. Significance Level 0.00000000

Q (16-0) = 268.7718. Significance Level 0.00000000

Q (20-0) = 275.1352. Significance Level 0.00000000

Q (24-0) = 275.2292. Significance Level 0.00000000

Cependant, bien que la saisonnalité ait amélioré sensiblement notre  $R^2$  explicatif, elle n'a pas permis une amélioration de la qualité des résidus en termes de stationnarité. Aussi, afin de ne pas alourdir notre modèle, nous allons continuer à étudier notre variable sans pour autant la dessaisonnaliser.

### 5.2.2 Estimation du modèle autorégressif

Comme nous l'avons spécifié plus haut pour cet exemple, nous allons nous intéresser au cas d'un modèle autorégressif de premier ordre.

Linear Regression - Estimation by Least Squares

Dependent Variable RESIDU

Monthly Data from 2004:01 To 2009:12

Usable Observations 72 Degrees of Freedom 70

Centered R\*\*2 0.926531 R Bar \*\*2 0.925482

Uncentered R\*\*2 0.926551 T x R\*\*2 66.712

Mean of Dependent Variable 0.221680242

Std Error of Dependent Variable 13.587645266

Standard Error of Estimate 3.709162288

Sum of Squared Residuals 963.05194167

Regression F (1, 70) 882.7854

Significance Level of F 0.000000000

Durbin-Watson Statistic 0.465900

Variable	Coeff	Std Error	T-Stat	Signif
*****				
1. Constant	0.0157614395	0.4371839054	0.03605	0.97134335
2. RESIDU {1}	0.9619255436	0.0323753070	29.71170	0.00000000

On peut observer de cette sortie qu'une régression de la variable résidu sur elle-même décalée d'une période donne des résultats significatifs. On obtient donc un modèle plutôt robuste où la dernière valeur de la variable résidu a un pouvoir explicatif considérable.

## 5.3 Prédiction du modèle autorégressif

Afin de pouvoir effectuer notre prédiction nous aurons besoin de pallier à deux problèmes identifiés précédemment pour rendre notre série stationnaire :

- On a établi précédemment qu'il existe une composante tendancielle dans notre variable d'intérêt à savoir le WPU10. Cette composante est représentée par la droite de tendance d'équation :  $y = 0,866 * x + 138,668$  tel que notre X est une variable correspondant aux dates de publications du WPU10. Ainsi si l'on veut prédire WPU10 il nous faudra prédire la composante du WPU10 non prédite par la tendance, qui se retrouve dans l'erreur d'estimation du WPU10 lorsqu'on l'estime par régression tendancielle. Donc pour continuer nos analyses pour effectuer une prédiction on se focalisera sur cette deuxième composante. Pour ce faire, une manière simple de se débarrasser de la composante tendance est de prendre le Log de notre série comme variable d'étude.
- De même il nous faudra faire la différenciation de notre variable Log nouvellement créée pour accentuer cette stationnarité.

### 5.3.1 Calibrage du modèle autorégressif d'ordre 1

Tableau 5-1 Régression de la première différence du log de WPU10

Input: WPU10 (Memoire historical data) Transformations: ln(x),D(1) Model:(0,1,1) MS Residual= .00019						
Paramet.	Param.	Asympt. Std.Err.	Asympt. t( 70)	p	Lower 95% Conf	Upper 95% Conf
Constant	0,.005714	0,.002793	2,.0456	0,.044555	0,.000143	0,.011284
q(1)	-0,.738333	0,.071095	-10,.3852	0,.000000	-0,.880127	-0,.596539

On constate de la sortie ci-dessus qu'il existe une autocorrélation. Bien que celle-ci ne soit pas dans la variable elle-même mais dans ses résidus tel que l'on a établi précédemment. On notera que le paramètre des résidus dans ce modèle-ci diffère du précédent puisque l'on n'a pas affaire aux mêmes résidus. Les premiers étant issus de la régression tendancielle du WPU10 alors que ceux-ci sont issus du calibrage d'une nouvelle variable correspondant à la première différence du log de WPU10. Cette autocorrélation a toujours un lag d'une période.

### 5.3.2 Test historique du modèle

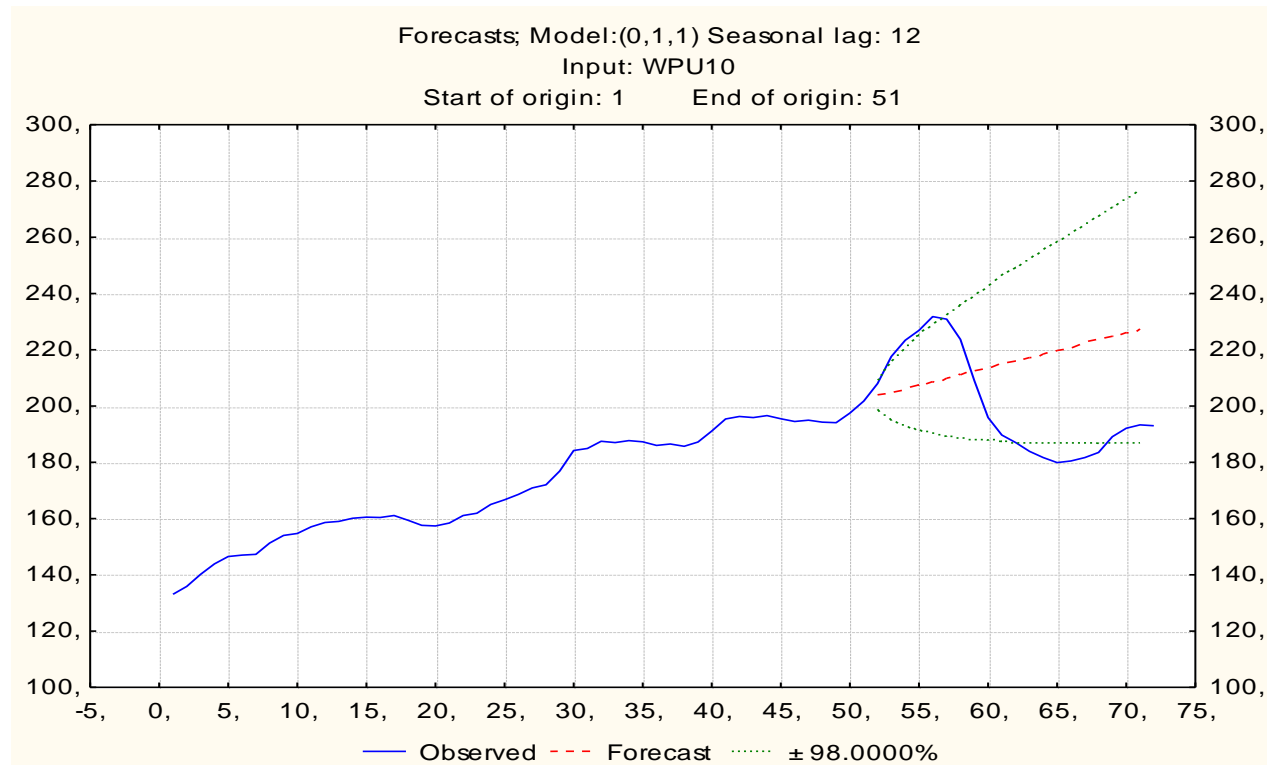


Figure 5-6 Test historique du WPU10 selon un modèle autorégressif

On a pris le modèle calibré à la section (5.3.1) et on a fait une prévision à l'intérieur de l'échantillon qui nous a servi à établir le modèle. Il apparaît clairement que le modèle autorégressif semble suivre la variable réelle aussi longtemps qu'il n'y a pas de changement de tendance. À cet égard sur notre période de test de janvier 2008 à décembre 2009, nous obtenons une corrélation négative de 0,7, une erreur moyenne de 26,7 ayant un écart type de 24,4. Ainsi sur la base du test historique, lorsqu'il y a un changement de régime ce type de modèle n'est pas efficace.

### 5.3.3 Prévision sans changement de régime avec une partie hors échantillon

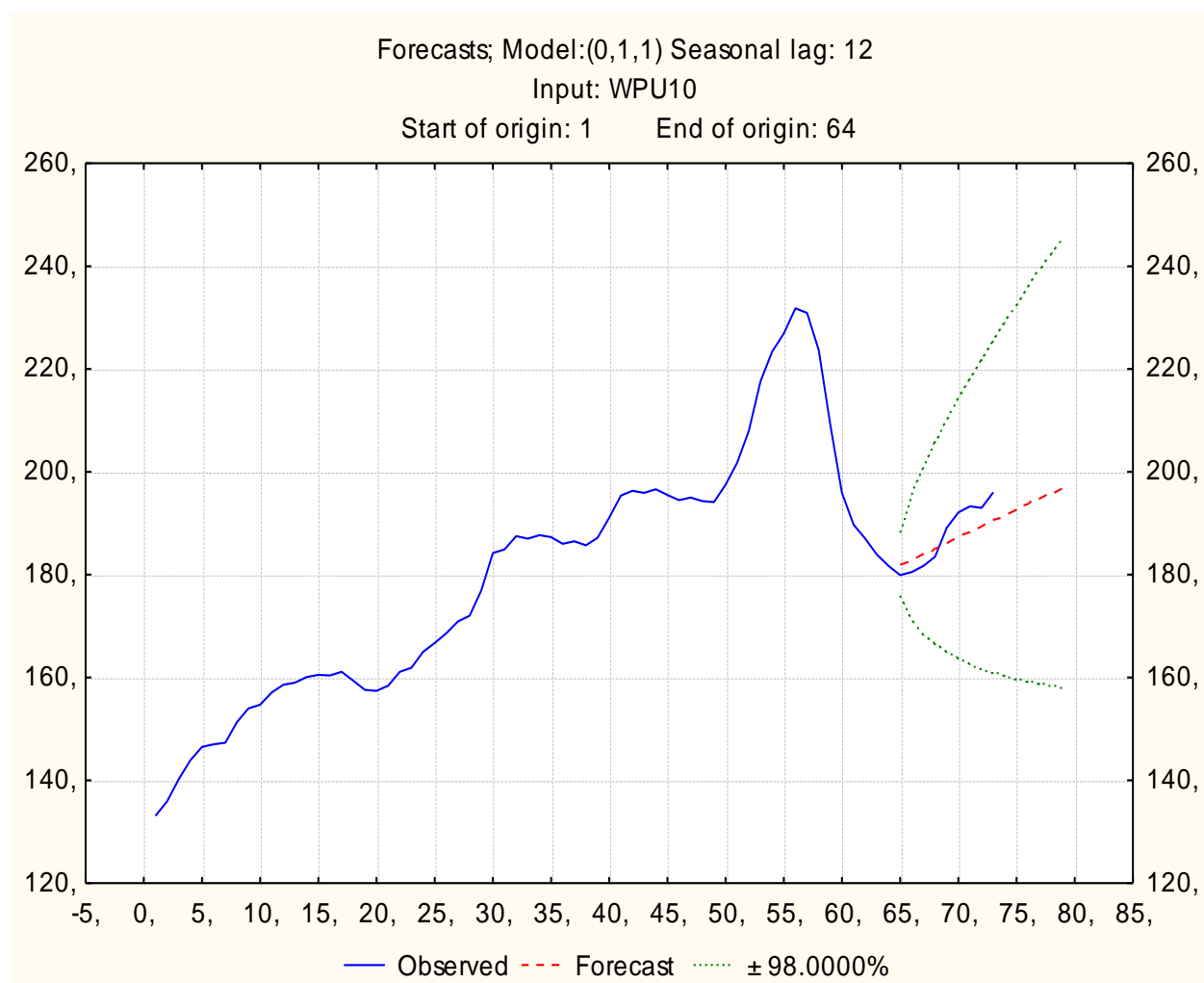


Figure 5-7 Prévision du WPU10 selon un modèle autorégressif

Pour une prévision d'avril 2009 à décembre 2009 avec en plus une prévision hors échantillon de janvier 2010 à juin 2010, on note que du fait de l'absence d'un changement de régime notre modèle est plus robuste avec des indicateurs très performants : corrélation (97,9%), moyenne des erreurs (5,37) et écart type des erreurs (5,88). Il semblerait donc que ce deuxième modèle est plus approprié pour capturer une tendance à long terme dans un environnement peu changeant.



## CONCLUSION

Nous avons étudié un indice, le WPU10, publié par le BLS qui est un agrégat des prix des métaux qui entrent dans la production industrielle.

À cet égard, on a étudié dans une première partie la composition de la variance du WPU10, puis nous avons étudié l'impact des indices sous-composants du WPU10 sur sa variance. Ainsi, après avoir établi que le modèle régressant la relation du WPU10 et ses sous-composants était linéaire, l'analyse de sensibilité de la variance du WPU10 par rapport à ses sous-composants grâce à l'indice SRC ou « Standardized Regression Coefficient » nous a révélé que moins de la moitié des indices le composant participaient pour une grande majorité à sa variance. Aussi, si l'on voulait étudier plus amplement le WPU10 et le prédire il nous faudrait étudier ces indices sous composants influents et leurs variances.

Tableau 5-2 Sous-composants influents du WPU10 et les matières premières qui les impactent

	Prix des matières premières:	Aluminium	Aluminium-moyenne mobile 6 mois	Cuivre	Cuivre-moyenne mobile 6 mois	Etain	Acier
Sous composants du WPU10	Description						
101211	Carbon steel scrap						X
10170301	Hot rolled sheet and strip, incl. tin mill						X
10170401	Hot rolled bars, plates, and structural shape			X		X	X
10170601	Steel pipe and tube		X				X
10170701	Cold rolled sheets and strip	X					X
1021	Nonferrous metal ores			X			
102201	Primary nonferrous metals			X			
102502	Copper and brass mill shapes			X			
1026	Nonferrous wire and cable			X			
1074	Struct., arch., pre-eng. metal products	X			X		X

L'étude de ces indices sous-composants du WPU10 nous a permis d'établir que dans leur ensemble ils sont sensibles au prix de quatre matières premières : le cuivre, l'aluminium, l'étain et l'acier. On a ainsi pu établir une relation directe entre les variations du WPU10 qui est un

indice et les variations des prix de métaux que l'on connaît mieux et pour lesquels grâce au marché financier nous disposons d'une manière quasi continue de l'information. Pour arriver à un tel résultat nous avons utilisé autant les méthodes de criblage pour identifier les quatre matières premières parmi l'ensemble des matières premières présentes sur le marché, que les méthodes traditionnelles d'analyse de régression et d'analyse de variance.

Après avoir établi que le WPU10 était sensible au cuivre, à l'aluminium, à l'étain et à l'acier, nous avons voulu étudier la robustesse de cette sensibilité que l'on venait d'établir. Aussi, nous nous sommes lancés dans un exercice de prévision du WPU10. Pour ce faire, on a utilisé le modèle de diffusion utilisé en analyse financière pour effectuer une prévision future des prix des matières premières pour en déduire les valeurs futures du WPU10. Le choix du modèle de diffusion provient du fait qu'à l'instar de l'analyse de sensibilité locale qui calcule des dérivées autour d'une valeur connue pour évaluer une sensibilité, le modèle de diffusion également nous permet de calculer un ensemble de solutions autour d'une valeur locale connue soit la valeur espérée du WPU10 tirée des prix futurs des matières premières. Il est ressorti de cet exercice de prévision que le modèle de diffusion comme outil de prévision est extrêmement efficace en ce sens qu'il permet de capturer dans une grande mesure les différents scénarios possibles dans le futur en contrôlant la variabilité de la variable que l'on cherche à comprendre ou expliquer. Le modèle basé sur la simulation de Monte Carlo s'est révélé un outil précieux à la réalisation de nos multiples simulations, nous permettant d'appliquer des chocs spécifiques à certains endroits dans le temps, nous offrant ainsi l'opportunité de créer des événements futurs à travers le marché financier pour en observer les effets sur les prix des matières premières et sur l'indice WPU10.

Enfin, dans la mesure où l'indice WPU10 est une série chronologique, nous avons cherché à savoir si la prévision du WPU10 pouvait se faire selon des méthodes d'analyse de séries temporelles qui ont la spécificité de ne pas faire appel à la variance de la variable dans la prévision contrairement au modèle de diffusion.

Tableau 5-3 Tableau comparatif des deux modèles de prévision

Indice synthétique	Corrélation avec l'indice réel élevée	Erreur De prévision faible	Prévision dans un régime stable	Prévision dans un régime changeant
Modèle de diffusion	✓	✓	✓	✓
Modèle auto régressif	✗	✗	✓	✗

Il est ressorti de la comparaison avec le modèle autorégressif d'ordre 1 que notre modèle principal est plus efficace tant dans sa flexibilité que dans la précision de ses prédictions, ceci autant dans un contexte historique basé sur l'échantillon de base que l'on a utilisé pour construire nos modèles que dans un contexte de prévision hors échantillon.

Ainsi, après avoir présenté les différents méthodes et indices de sensibilité et une application de l'analyse de sensibilité dans un contexte de prévision, nous pensons qu'une suite possible pour ce mémoire serait le développement d'un modèle de type VAR ou vecteur autorégressif pour voir s'il est plus efficace que notre modèle. En effet, le modèle autorégressif se base surtout sur l'indice principal WPU10 comme variable, alors que le modèle VAR à l'instar de notre modèle principal pourrait essayer de construire un indice synthétique en créant des sous modèles autorégressifs pour les sous-composants de l'indice WPU10 et chercher ainsi à reconstruire un nouvel indice synthétique WPU10 que l'on pourrait comparer à celui que l'on a développé par analyse de sensibilité.

## BIBLIOGRAPHIE

1. Antoniadis, A. (2003), Analyse de la variance fonctionnelle pour l'analyse de sensibilité globale, Groupement MOMAS (Modélisations Mathématiques et Simulations Numériques liées aux problèmes de gestion des déchets nucléaire).
2. Campolongo, F. (2005), Screening methods in sensitivity analysis: The Morris method and its applications, JRC-Ispra (Joint research centre, European commission).
3. François, P. (2005), Les Produits dérivés financiers, Dunod.
4. Golda, D., Lagnemma, K., Dubowsky, S. (2008), Probabilistic Modelling and Analysis of High-Speed Rough-Terrain Mobile Robots, Inédit Massachusetts Institute of Technology, USA, consulté le 10/07/2009, tiré de [http://robots.mit.edu/publications/papers/2004\\_04\\_Gol\\_Iag\\_Dub.pdf](http://robots.mit.edu/publications/papers/2004_04_Gol_Iag_Dub.pdf)
5. Iooss, B., revue sur l'analyse de sensibilité globale des modèles numériques.
6. Jacques, J. (2005), Contributions à l'analyse de sensibilité et à l'analyse discriminante généralisée, Université Joseph Fourier - Grenoble 1.
7. John C. Hull (1997). Options Futures and Other Derivatives, third edition, Prentice Hall, Upper Saddle River, NJ.
8. Kaya, H., & Kaplan, M., & Saygın, H., A recursive algorithm for finding HDMR terms for sensitivity analysis.

9. Nadia, P., IOOSS, B., (2008), Analyse de sensibilité de modèle : Calibration, Centre d'Étude de Cadarache, Direction de l'Énergie Nucléaire, France.
  
10. Paskov, S., Traub, J. (1995), Faster valuation of financial derivatives, inédit department of computer science, Columbia university, USA. Consulté le 01/09/2009, Tiré de [www.cs.columbia.edu/techreports/cucs-030-96.ps.gz](http://www.cs.columbia.edu/techreports/cucs-030-96.ps.gz)
  
11. Ratto, M., State dependent parameter modelling, Meta modelling and sensitivity analysis, IPSC-JSC, European commission, Consulté le 20/aout/2009, tiré de <http://www.jrc.cec.eu.int/uasa>
  
12. Redl, C., Haas. R., Huber, C., Böhm, B. (2008), Price formation in electricity forward markets and the relevance of systematic forecast errors, Energy Economics, 31(5), 356-364
  
13. Rombouts, J. (2006), Techniques d'analyse et de prévision de la conjoncture, Inédit HEC Montréal, Québec, Canada.
  
14. Saltelli, A., Chan, K. , and Scott, E.M., editors. Sensitivity analysis. Wiley Series in Probability and Statistics. Wiley, 2000.
  
15. Saltelli, A. (2006), The critique of modelling and sensitivity analysis in the scientific discourse, an overview of good practices. Institute for the protection and security of the citizen, European commission Joint research centre.
  
16. Sobol, M. (1993), Sensitivity estimates for nonlinear mathematical models, Mathematical Modelling and Computational Experiments.

17. Tarantola, S. (2006), Variance based methods for sensitivity analysis, European commission joint research centre of ISPRA, Venice, Italy.
18. Turyani, T. et al.(2005), Similarity of local sensitivity functions, Institute of Chemistry, Eötvös University (ELTE), Budapest, Hungary.
19. <http://www.bls.gov>

## ANNEXE 1 – Feuille Excel d'initialisation de la simulation

### I. Initialisation de la simulation

Entrer votre percentile pour le range initial

Range		
Bas		Haut
0,16	vs	0,40

Entrer votre percentile pour le range secondaire

0,050	vs	0,60
-------	----	------

### II. Entrer votre matrice des corrélations pour avoir la décomposition de Cholesky

Commodités	MATRICE CORRELATION				
	Copper	Al	TIN	HRC	
Copper	1		0,44	0,33	0,2
Al	0,44	1		0,45	0,56
TIN	0,33	0,45	1		0,34
HRC	0,2	0,56	0,34	1	

Commodités	Matrice de Corrélation selon la décomposition de CHOLESKY				
	Copper	Al	TIN	HRC	SSS
Copper	1		0	0	0
Al	0,44	0,897997773		0	0
TIN	0,33	0,339421777	0,880847806		0
HRC	0,2	0,525613776	0,108526396	0,8197269	

**Expanding Loop pour qu'on puisse simuler pour la période voulue**

### III. Mettre à jour les données sur les commodités

Entrer le nombre de mois de prévision voulue (Le nombre de mois doit être inférieur ou égal à 12)	12
Entrer la taille de la simulation voulue (On conseil une simulation entre 10-20 si on veut)	50
Tester le modèle et dans les alentours de 100 pour avoir des résultats solides)	
Voir les données de la simulation?(True/False)	False

## Données prix commodities

Commodités	Unité	Date mis à jour	Source	Prix Actuel	3 Mois	15 Mois	27 mois
Copper	\$ USD/ton.	04-janv-10	LME	7464,00	7504,00	7555,00	7555,00
Aluminium	\$ USD/ton.	04-janv-10	LME	2247,00	2261,00	2388,00	2463,00
Tin	\$ USD/ton.	04-janv-10	LME	17450,00	17550,00	17585,00	17615,00
HRC	\$ USD/ton.	04-janv-10	NYMEX Steel Future	565,50	606,30	622,80	622,80

## Données croissance prix commodities

Commodités	Moyenne croissance mensuelle 3 mois	Moyenne croissance mensuelle 15 mois	Moyenne croissance mensuelle 27 mois
Copper	0,178%	0,056%	0,000%
Aluminium	0,207%	0,455%	0,258%
Tin	0,190%	0,017%	0,014%
HRC	2,322%	0,224%	0,000%

## Données volatilité commodities

Commodités	Volatilité actuelle	Volatilité 3 mois	Volatilité 15 mois	Volatilité 27 mois et plus
Copper	8,7%	8,7%	8,7%	8,7%
Aluminium	11,5%	11,5%	11,5%	11,5%
Tin	10,1%	11,0%	11,0%	11,0%
HRC	13,0%	13,0%	13,0%	13,0%

## Données Modèles

Commodités	% volatilité du WPU10
Copper	8,0%
Copper 6 MA	9,8%
Aluminium	6,2%
Aluminium 6 Ma	0,7%
Tin	2,1%
HRC	13,0%
Constante	60,3%



## ANNEXE 2 – Statistiques des nombres aléatoires multivariés

Moyenne	Month1	Month2	Month3	Month4	Month5	Month6
<b>Copper</b>	0,02531058	-0,38375839	0,15112265	0,23291544	0,17151572	0,18838942
<b>Aluminium</b>	0,0222729	-0,27772845	0,11551727	0,00230972	-0,02488918	0,07237629
<b>TIN</b>	0,39310716	0,03835494	0,25344027	0,09654235	0,20633796	0,11853246
<b>HRC</b>	0,06800596	-0,22504967	0,24825059	0,02858665	-0,22626859	0,11244811

Stat Percentile 0,05	Month1	Month2	Month3	Month4	Month5	Month6
<b>Copper</b>	-1,57287439	-1,70318557	-1,03485957	-1,57462324	-0,50800712	-1,0620997
<b>Aluminium</b>	-1,12170645	-1,58117093	-0,99705042	-1,47308296	-1,54077279	-1,07073344
<b>TIN</b>	-0,42935319	-1,79967709	-1,74064428	-1,51151166	-1,78465527	-1,18624387
<b>HRC</b>	-1,40391389	-1,53846935	-1,29269513	-1,48473642	-1,43691641	-0,97359825

Stat Percentile 0,16	Month1	Month2	Month3	Month4	Month5	Month6
<b>Copper</b>	-0,96353181	-1,38450293	-0,44277794	-0,96016894	-0,35130324	-0,57975452
<b>Aluminium</b>	-0,61193478	-0,8856958	-0,6241884	-0,92211172	-0,84934	-0,9301905
<b>TIN</b>	-0,32139435	-1,17279767	-0,74304902	-0,64426058	-1,04098935	-0,68045605
<b>HRC</b>	-0,71909749	-0,95380044	-1,059264	-1,14889632	-1,18849547	-0,75215424

Stat Percentile 0,4	Month1	Month2	Month3	Month4	Month5	Month6
<b>Copper</b>	-0,09078724	-0,52757193	0,0290525	-0,0586095	-0,13821606	-0,19736353
<b>Aluminium</b>	-0,08365608	-0,18111493	-0,27588602	-0,11445893	-0,03398175	-0,22874304
<b>TIN</b>	0,2252302	-0,31248163	0,28803098	0,08313704	0,26022043	0,00766477
<b>HRC</b>	-0,0110504	-0,38559801	-0,01269021	-0,18330171	-0,83156464	-0,2901082

Stat Percentile 0,8	Month1	Month2	Month3	Month4	Month5	Month6
<b>Copper</b>	1,01303782	0,50758382	0,70265329	1,12721826	0,73989442	1,02694291
<b>Aluminium</b>	0,69270372	0,2575201	0,94071585	0,92681771	0,44318313	0,83332545
<b>TIN</b>	0,99613768	0,85283286	0,80728392	0,86530444	1,10017398	0,58179689
<b>HRC</b>	0,74223069	0,63085642	1,17732153	0,93983059	0,72591076	0,60548429

### ANNEXE 3 – Volatilité et taux de croissance mensuel par scénario

Monthly Volatility	Month1	Month2	Month3	Month4	Month5	Month6
Copper	0,08660254	0,08660254	0,08660254	0,08660254	0,08660254	0,08660254
Aluminium	0,11547005	0,11547005	0,11547005	0,11547005	0,11547005	0,11547005
TIN	0,10969655	0,10969655	0,10969655	0,10969655	0,10969655	0,10969655
HRC	0,12990381	0,12990381	0,12990381	0,12990381	0,12990381	0,12990381

Avg.Monthly growth rate	Month1	Month2	Month3	Month4	Month5	Month6
Copper	0,00178158	0,00178158	0,00056445	0,00056445	0,00056445	0,00056445
Aluminium	0,0020704	0,0020704	0,00455408	0,00455408	0,00455408	0,00455408
TIN	0,00190477	0,00190477	0,00016603	0,00016603	0,00016603	0,00016603
HRC	0,02322154	0,02322154	0,00223754	0,00223754	0,00223754	0,00223754

Expected Growth Rate	Month1	Month2	Month3	Month4	Month5	Month6
Copper	0,00397354	-0,03145287	0,01365205	0,02073552	0,01541815	0,01687945
Aluminium	0,00464225	-0,02999892	0,01789287	0,00482079	0,00168013	0,01291138
TIN	0,04502727	0,00611217	0,02796755	0,01075639	0,02280059	0,01316863
HRC	0,03205577	-0,00601327	0,03448624	0,00595106	-0,02715561	0,01684498

Growth Rate Percentile 0,05	Month1	Month2	Month3	Month4	Month5	Month6
Copper	-0,13443334	-0,14571861	-0,08905702	-0,13580192	-0,04343026	-0,09141608
Aluminium	-0,1274531	-0,18050749	-0,11057538	-0,16554288	-0,17335903	-0,11908356
TIN	-0,0451938	-0,1955136	-0,19077665	-0,16564159	-0,1956045	-0,12996084
HRC	-0,15915222	-0,17663149	-0,16568848	-0,19063538	-0,18442337	-0,12423658

Growth Rate Percentile 0,16	Month1	Month2	Month3	Month4	Month5	Month6
Copper	-0,08166272	-0,11811989	-0,03778125	-0,08258862	-0,0298593	-0,04964377
Aluminium	-0,06858974	-0,10020094	-0,06752098	-0,10192221	-0,09351925	-0,10285506
TIN	-0,03335108	-0,12674709	-0,08134389	-0,07050714	-0,11402691	-0,07447766
HRC	-0,07019196	-0,10068077	-0,13536489	-0,14700847	-0,15215255	-0,09547016

Growth Rate Percentile 0,4	Month1	Month2	Month3	Month4	Month5	Month6
Copper	-0,00608082	-0,04390749	0,00308047	-0,00451128	-0,01140541	-0,01652773
Aluminium	-0,00758937	-0,01884295	-0,02730249	-0,00866249	0,00063021	-0,02185889
TIN	0,02661174	-0,03237339	0,03176203	0,00928587	0,02871131	0,00100683
HRC	0,02178605	-0,02686911	0,00058904	-0,02157405	-0,10578587	-0,03544862

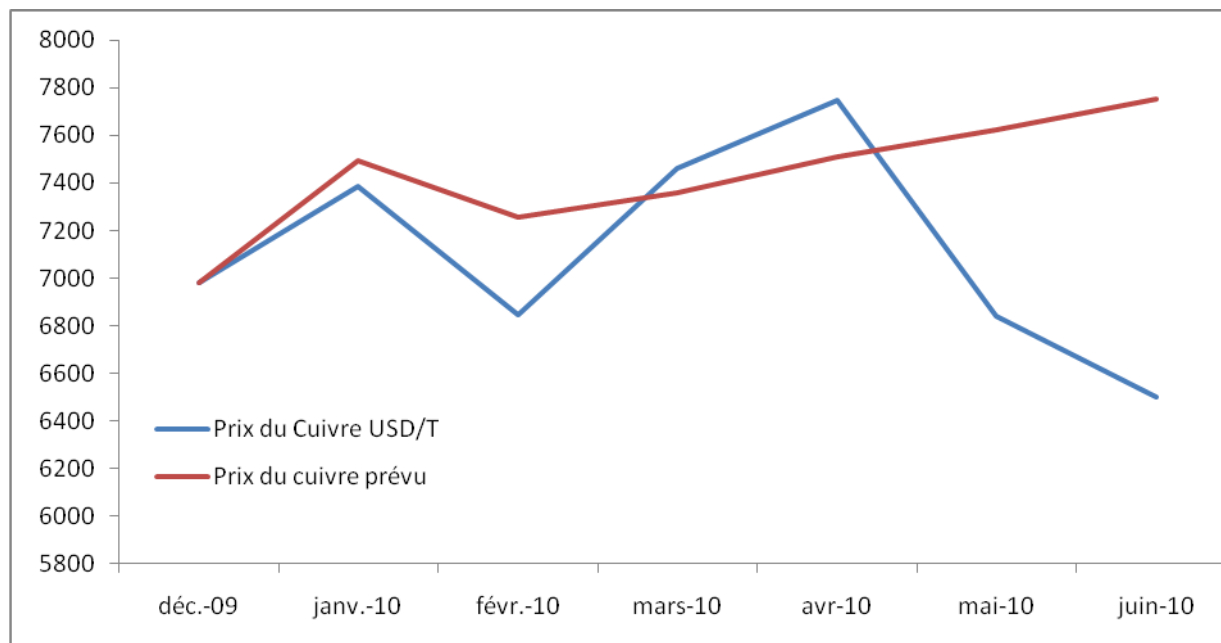
  

Growth Rate Percentile 0,8	Month1	Month2	Month3	Month4	Month5	Month6
Copper	0,08951323	0,04573963	0,06141601	0,09818441	0,06464119	0,08950031
Aluminium	0,08205694	0,03180626	0,11317859	0,11157377	0,05572846	0,10077822
TIN	0,11117764	0,09545759	0,08872229	0,09508694	0,12085132	0,06398714
HRC	0,11964013	0,10517219	0,1551761	0,12432512	0,09653612	0,08089226

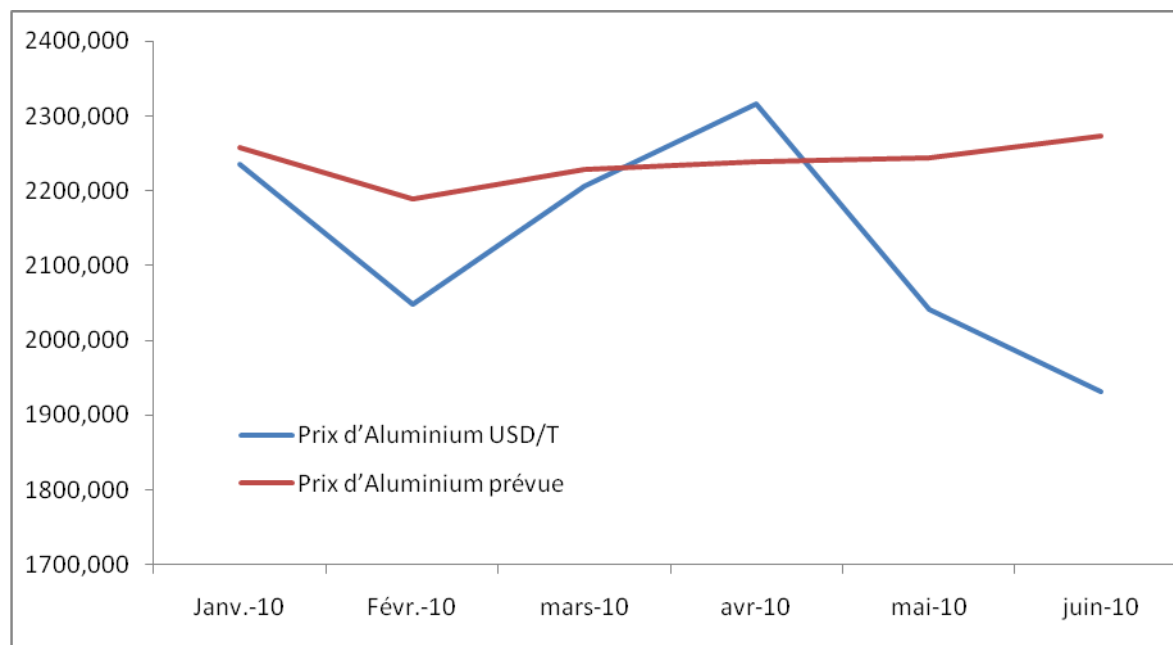
## ANNEXE 4 – Prévision

Forecast Expected	Prix Actuel	Month1	Month2	Month3	Month4	Month5	Month6
Copper	7464,00	7493,66	7257,96	7357,05	7509,60	7625,38	7754,10
MovingAvg. Copper	7464,00	7493,66	7375,81	7369,56	7404,57	7448,73	7499,62
Aluminium	2247,00	2257,43	2189,71	2228,89	2239,64	2243,40	2272,36
MovingAvg. Aluminium	2247,00	2257,43	2223,57	2225,34	2228,92	2231,81	2238,57
TIN	17450,00	18235,73	18347,19	18860,31	19063,18	19497,83	19754,59
HRC	565,50	583,63	580,12	600,12	603,70	587,30	597,19
Forecast Percentile 0,05	Prix Actuel	Month1	Month2	Month3	Month4	Month5	Month6
Copper	7464,00	6460,59	5519,16	5027,64	4344,88	4156,18	3776,24
MovingAvg. Copper	7464,00	6460,59	5989,88	5669,13	5338,07	5101,69	4880,78
Aluminium	2247,00	1960,61	1606,71	1429,05	1192,48	985,75	868,36
MovingAvg. Aluminium	2247,00	1960,61	1783,66	1665,46	1547,21	1434,92	1340,49
TIN	17450,00	16661,37	13403,84	10846,70	9050,04	7279,81	6333,72
HRC	565,50	475,50	391,51	326,64	264,37	215,62	188,83
Forecast Percentile 0,16	Prix Actuel	Month1	Month2	Month3	Month4	Month5	Month6
Copper	7464,00	6854,47	6044,82	5816,44	5336,07	5176,74	4919,74
MovingAvg. Copper	7464,00	6854,47	6449,64	6238,58	6012,95	5845,71	5691,38
Aluminium	2247,00	2092,88	1883,17	1756,02	1577,04	1429,56	1282,52
MovingAvg. Aluminium	2247,00	2092,88	1988,02	1910,69	1827,28	1747,73	1670,20
TIN	17450,00	16868,02	14730,05	13531,85	12577,76	11143,56	10313,61
HRC	565,50	525,81	472,87	408,86	348,75	295,69	267,46
Forecast Percentile 0,4	Prix Actuel	Month1	Month2	Month3	Month4	Month5	Month6
Copper	7464,00	7418,61	7092,88	7114,73	7082,63	7001,85	6886,13
MovingAvg. Copper	7464,00	7418,61	7255,75	7208,74	7177,21	7142,14	7099,47
Aluminium	2247,00	2229,95	2187,93	2128,19	2109,76	2111,09	2064,94
MovingAvg. Aluminium	2247,00	2229,95	2208,94	2182,02	2163,96	2153,38	2138,64
TIN	17450,00	17914,37	17334,43	17885,00	18051,08	18569,35	18588,05
HRC	565,50	577,82	562,29	562,63	550,49	492,25	474,80
Forecast Percentile 0,8	Prix Actuel	Month1	Month2	Month3	Month4	Month5	Month6
Copper	7464,00	8132,13	8504,09	9026,37	9912,62	10553,39	11497,92
MovingAvg. Copper	7464,00	8132,13	8318,11	8554,20	8893,80	9225,72	9604,42
Aluminium	2247,00	2431,38	2508,72	2792,65	3104,23	3277,23	3607,50
MovingAvg. Aluminium	2247,00	2431,38	2470,05	2577,58	2709,24	2822,84	2953,62
TIN	17450,00	19390,05	21240,98	23125,53	25324,46	28384,96	30201,23
HRC	565,50	633,16	699,75	808,33	908,83	996,56	1077,18

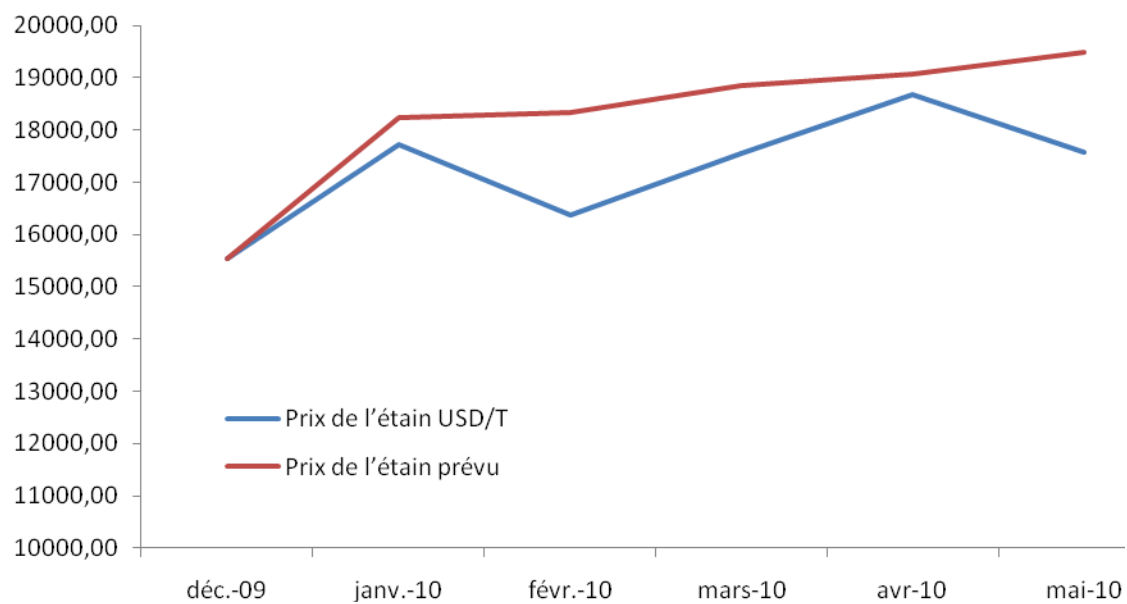
## ANNEXE 5 – Détails de la prévision



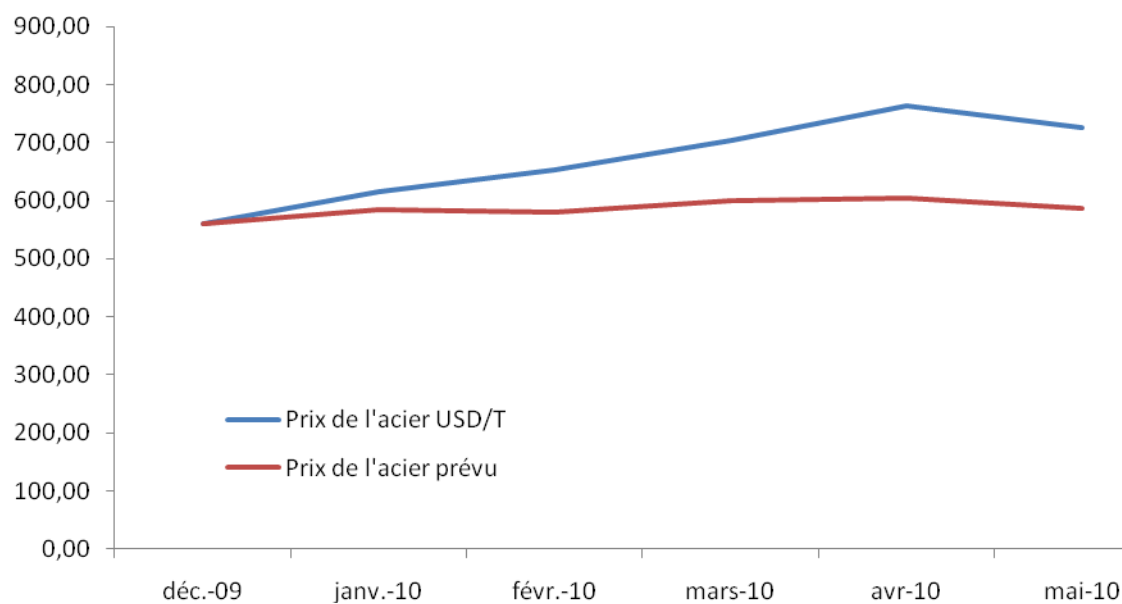
Prévision du Cuivre	déc.-09	janv.-10	févr.-10	mars-10	avril-10	mai-10	juin-10
Prix du Cuivre USD/T	6981,710	7386,250	6848,180	7462,830	7745,080	6837,680	6499,300
Prix du cuivre prévu	6981,710	7493,659	7257,961	7357,048	7509,600	7625,384	7754,096
Différence	0,000	107,409	409,781	-105,782	-235,480	787,704	1254,796



Prévision dz l'Aluminium	Janv.-10	Févr.-10	mars-10	avr-10	mai-10	juin-10
Prix d'Aluminium USD/T	2235,150	2048,930	2205,830	2316,720	2040,530	1931,390
Prix d'Aluminium prévue	2257,431	2189,711	2228,891	2239,636	2243,399	2272,364
Différence	22,281	22,281	22,281	22,281	22,281	22,281



Prévision de l'étain	déc.-09	janv.-10	févr.-10	mars-10	avr-10	mai-10	juin-10
Prix de l'étain USD/T	15546,90	17714,75	16361,75	17549,35	18683,50	17566,05	17319,77
Prix de l'étain prévu	15546,90	18235,73	18347,19	18860,31	19063,18	19497,83	19754,59
Différence	0,00	520,98	1985,44	1310,96	379,68	1931,78	2434,82



Prévision de l'acier	déc.-09	janv.-10	févr.-10	mars-10	avr-10	mai-10	juin-10
Prix de l'acier USD/T	560,96	615,75	654,08	704,58	763,70	726,75	674,23
Prix de l'acier prévu	560,96	583,63	580,12	600,12	603,70	587,30	597,19
Différence	0,00	-32,12	-73,96	-104,46	-160,00	-139,45	-77,04

## ANNEXE 6 – Code VBA

```
Sub CodeFinal()
```

'Lorsqu'une note est précédé d'un guillemet c'est qu'une nouvelle partie du code commence

```
***** Genere Nombre aleatoire Multivarié *****
```

```
' On crée les feuilles par mois de simulation et calcul les
```

```
' Variables aléatoires distribués normalement mais corréllées
```

```
' Macro enregistrée le 12/02/2010 par A.M.Abbassi
```

```
Dim TailleSim As Long
```

```
Dim Z As Long ' Nombre de mois qui constitue notre forecast
```

```
Z = Sheets("Initialisation").Range("D32")
```

```
TailleSim = Sheets("Initialisation").Range("D34")
```

```
'On crée les tableaux dans les différentes feuilles
```

```
For i = 1 To Z
```

```
    Sheets.Add.Move After:=Sheets(Sheets.Count)
```

```
    Sheets(Sheets.Count).Cells(1, 1) = "Simulation"
```

```
    Sheets(Sheets.Count).Cells(1, 2) = "Copper"
```

```
    Sheets(Sheets.Count).Cells(1, 3) = "Aluminium"
```

```
Sheets(Sheets.Count).Cells(1, 4) = "TIN"
```

```
Sheets(Sheets.Count).Cells(1, 5) = "HRC"
```

```
For j = 1 To TailleSim
```

```
  Sheets(Sheets.Count).Cells(j + 1, 1) = j
```

```
Next j
```

'On génère des nombres aléatoires

```
Sheets.Add.Move After:=Sheets(Sheets.Count - 1)
```

```
Dim Num As Integer
```

```
Num = i * 456 + 1000 / i
```

```
Randomize (Num)
```

```
Sheets(Sheets.Count - 1).Activate
```

```
ThisWorkbook.Worksheets(Sheets.Count - 1).Range("A1", "D" &  
TailleSim).FormulaLocal = "=loi.normale.inverse(alea(); 0; 1)"
```

```
ThisWorkbook.Worksheets(Sheets.Count - 1).Range("A1", "D" &  
TailleSim).CurrentRegion.Select
```

```
ActiveWorkbook.Names.Add Name:="Tablo" & i, RefersTo:=Selection
```

'On selectionne le tableau précédemment définit dans notre dernière feuille

```
Sheets(Sheets.Count).Activate
```

```
Sheets(Sheets.Count).Range("B2", "E" & TailleSim + 1).Select
```

```
Selection.FormulaArray = "=MMULT(Tablo" & i &  
",Transpose(Initialisation!B23:E26))"
```



```
Application.ScreenUpdating = False
```

```
Sheets(Sheets.Count - 1).Name = "NbrAlea" & i
```

```
Sheets(Sheets.Count).Name = "Alea Month" & i
```

```
Next i
```

```
***** On construit les statistiques des nombres aleatoires*****'
```

```
' On Construit le tableau pour la Moyenne avec ses caractéristiques
```

```
ActiveWorkbook.Worksheets.Add.Move After:=Sheets(Sheets.Count)
```

```
Sheets(Sheets.Count).Name = "Statistiques NbrAlea"
```

```
For i = 2 To 5
```

```
For h = 2 To Z + 1
```

```
For w = 0 To 4
```

```
Worksheets("Statistiques NbrAlea").Cells(1, 1) = "Moyenne"
```

```
Worksheets("Statistiques NbrAlea").Cells(11, 1) = "Stat Percentile " &  
Sheets("Initialisation").Range("E8")
```

```
Worksheets("Statistiques NbrAlea").Cells(21, 1) = "Stat Percentile " &  
Sheets("Initialisation").Range("E5")
```

```
Worksheets("Statistiques NbrAlea").Cells(31, 1) = "Stat Percentile " &  
Sheets("Initialisation").Range("G5")
```

```
Worksheets("Statistiques NbrAlea").Cells(41, 1) = "Stat Percentile " &  
Sheets("Initialisation").Range("G8")
```

Worksheets("Statistiques NbrAlea").Cells(1 + 10 \* w, 1).Interior.ColorIndex = 5

Worksheets("Statistiques NbrAlea").Cells(1 + 10 \* w, 1).Font.Bold = True

Worksheets("Statistiques NbrAlea").Cells(2 + 10 \* w, 1).Font.Bold = True

Worksheets("Statistiques NbrAlea").Cells(3 + 10 \* w, 1).Font.Bold = True

Worksheets("Statistiques NbrAlea").Cells(4 + 10 \* w, 1).Font.Bold = True

Worksheets("Statistiques NbrAlea").Cells(5 + 10 \* w, 1).Font.Bold = True

Worksheets("Statistiques NbrAlea").Cells(1 + 10 \* w, 1).Font.ColorIndex = 2

Worksheets("Statistiques NbrAlea").Cells(2 + 10 \* w, 1) = "Copper"

Worksheets("Statistiques NbrAlea").Cells(3 + 10 \* w, 1) = "Aluminium"

Worksheets("Statistiques NbrAlea").Cells(4 + 10 \* w, 1) = "TIN"

Worksheets("Statistiques NbrAlea").Cells(5 + 10 \* w, 1) = "HRC"

Sheets("Statistiques NbrAlea").Cells(1 + 10 \* w, h) = " Month" & h - 1

Sheets("Statistiques NbrAlea").Cells(1 + 10 \* w, h).Interior.ColorIndex = 5

Sheets("Statistiques NbrAlea").Cells(1 + 10 \* w, h).Font.ColorIndex = 2

Sheets("Statistiques NbrAlea").Cells(1 + 10 \* w, h).Font.Bold = True

Sheets("Statistiques NbrAlea").Cells(i + 10 \* w, h).Interior.ColorIndex = 16

Sheets("Statistiques NbrAlea").Cells(i + 10 \* w, h).Font.ColorIndex = 2

Next w

Next h

Next i

```
For h = 2 To Z + 1
```

```
    For i = 2 To 5
```

```
        ThisWorkbook.Worksheets("Statistiques NbrAlea").Cells(i, h).Value =
Application.WorksheetFunction.Average(Worksheets("Alea Month" & h - 1).Range("'" & Chr(64
+ i) & "'" & 2, "'" & Chr(64 + i) & "'" & TailleSim))
```

```
    Next i
```

```
Next h
```

```
For h = 2 To Z + 1
```

```
    For i = 2 To 5
```

```
        ThisWorkbook.Worksheets("Statistiques NbrAlea").Cells(i + 10, h).Value =
Application.WorksheetFunction.Percentile(Worksheets("Alea Month" & h - 1).Range("'" &
Chr(64 + i) & "'" & 2, "'" & Chr(64 + i) & "'" & TailleSim), Sheets("Initialisation").Range("E8"))
```

```
    Next i
```

```
Next h
```

```
For i = 2 To 5
```

```
    For h = 2 To Z + 1
```

```
        ThisWorkbook.Worksheets("Statistiques NbrAlea").Cells(i + 20, h).Value =
Application.WorksheetFunction.Percentile(Worksheets("Alea Month" & h - 1).Range("'" &
Chr(64 + i) & "'" & 2, "'" & Chr(64 + i) & "'" & TailleSim), Sheets("Initialisation").Range("E5"))
```

```
    Next h
```

```
Next i
```

```
For i = 2 To 5
```

```
For h = 2 To Z + 1
```

```
    ThisWorkbook.Worksheets("Statistiques NbrAlea").Cells(i + 30, h).Value =  
    Application.WorksheetFunction.Percentile(Worksheets("Alea Month" & h - 1).Range("" &  
    Chr(64 + i) & "" & 2, "" & Chr(64 + i) & "" & TailleSim), Sheets("Initialisation").Range("G5"))
```

```
Next h
```

```
Next i
```

```
For i = 2 To 5
```

```
    For h = 2 To Z + 1
```

```
        ThisWorkbook.Worksheets("Statistiques NbrAlea").Cells(i + 40, h).Value =  
        Application.WorksheetFunction.Percentile(Worksheets("Alea Month" & h - 1).Range("" &  
        Chr(64 + i) & "" & 2, "" & Chr(64 + i) & "" & TailleSim), Sheets("Initialisation").Range("G8"))
```

```
    Next h
```

```
Next i
```

```
Columns("A:A").EntireColumn.AutoFit
```

```
'***** On cache les feuilles portant les nombres aleatoires*****'
```

```
For i = 1 To Z
```

```
    Sheets("NbrAlea" & i).visible = False
```

```
    Sheets("Alea Month" & i).visible = False
```

```
Next i
```

'\*\*\*\*On calcul le taux de croissance mensuel et On Construit le tableau pour chaque statistique avec ses caractéristique\*\*\*\*\*'

```
ActiveWorkbook.Worksheets.Add.Move After:=Sheets(Sheets.Count)
```

```
Sheets(Sheets.Count).Name = "Recapitulatif"
```

```
Worksheets("Recapitulatif").Cells(11, 1) = "Avg.Monthly growth rate"
```

```
Worksheets("Recapitulatif").Cells(11, 1).Interior.ColorIndex = 5
```

```
Worksheets("Recapitulatif").Cells(11, 1).Font.Bold = True
```

```
Worksheets("Recapitulatif").Cells(12, 1).Font.Bold = True
```

```
Worksheets("Recapitulatif").Cells(13, 1).Font.Bold = True
```

```
Worksheets("Recapitulatif").Cells(14, 1).Font.Bold = True
```

```
Worksheets("Recapitulatif").Cells(15, 1).Font.Bold = True
```

```
Worksheets("Recapitulatif").Cells(11, 1).Font.ColorIndex = 2
```

```
Worksheets("Recapitulatif").Cells(12, 1) = "Copper"
```

```
Worksheets("Recapitulatif").Cells(13, 1) = "Aluminium"
```

```
Worksheets("Recapitulatif").Cells(14, 1) = "TIN"
```

```
Worksheets("Recapitulatif").Cells(15, 1) = "HRC"
```

For i = 2 To 5

For h = 2 To Z + 1

Sheets("Recapitulatif").Cells(11, h) = " Month" & h - 1

Sheets("Recapitulatif").Cells(11, h).Interior.ColorIndex = 5

Sheets("Recapitulatif").Cells(11, h).Font.ColorIndex = 2

Sheets("Recapitulatif").Cells(11, h).Font.Bold = True

If h <= 3 Then

Sheets("Recapitulatif").Cells(i + 10, h) = Worksheets("Initialisation").Cells(50 + i, 2)

ElseIf h > 3 And h <= 15 Then

Sheets("Recapitulatif").Cells(i + 10, h) = Worksheets("Initialisation").Cells(50 + i, 3)

ElseIf h > 15 And h <= 27 Then

Sheets("Recapitulatif").Cells(i + 10, h) = Worksheets("Initialisation").Cells(50 + i, 4)

End If

Sheets("Recapitulatif").Cells(i + 10, h).Interior.ColorIndex = 16

Sheets("Recapitulatif").Cells(i + 10, h).Font.ColorIndex = 2

Next h

Next i

'On calcul la volatilité mensuelle

Worksheets("Recapitulatif").Cells(1, 1) = "Monthly Volatility"

Worksheets("Recapitulatif").Cells(1, 1).Interior.ColorIndex = 5

Worksheets("Recapitulatif").Cells(1, 1).Font.Bold = True

Worksheets("Recapitulatif").Cells(2, 1).Font.Bold = True

Worksheets("Recapitulatif").Cells(3, 1).Font.Bold = True

Worksheets("Recapitulatif").Cells(4, 1).Font.Bold = True

Worksheets("Recapitulatif").Cells(5, 1).Font.Bold = True

Worksheets("Recapitulatif").Cells(1, 1).Font.ColorIndex = 2

Worksheets("Recapitulatif").Cells(2, 1) = "Copper"

Worksheets("Recapitulatif").Cells(3, 1) = "Aluminium"

Worksheets("Recapitulatif").Cells(4, 1) = "TIN"

Worksheets("Recapitulatif").Cells(5, 1) = "HRC"

For i = 2 To 5

For h = 2 To Z + 1

Sheets("Recapitulatif").Cells(1, h) = " Month" & h - 1

Sheets("Recapitulatif").Cells(1, h).Interior.ColorIndex = 5

Sheets("Recapitulatif").Cells(1, h).Font.ColorIndex = 2

Sheets("Recapitulatif").Cells(1, h).Font.Bold = True

If h <= 3 And Worksheets("Initialisation").Cells(60 + i, 2) >=

Worksheets("Initialisation").Cells(60 + i, 3) Then

Sheets("Recapitulatif").Cells(i, h) = Worksheets("Initialisation").Cells(60 + i, 2)

```
Else: Sheets("Recapitulatif").Cells(i, h) = Worksheets("Initialisation").Cells(60 + i, 3)
```

```
End If
```

```
If h > 3 And h <= 15 Then
```

```
    Sheets("Recapitulatif").Cells(i, h) = Worksheets("Initialisation").Cells(60 + i, 3)
```

```
ElseIf h > 15 And h <= 27 Then
```

```
    Sheets("Recapitulatif").Cells(i, h) = Worksheets("Initialisation").Cells(60 + i, 4)
```

```
End If
```

```
    Sheets("Recapitulatif").Cells(i, h).Interior.ColorIndex = 16
```

```
    Sheets("Recapitulatif").Cells(i, h).Font.ColorIndex = 2
```

```
Next h
```

```
Next i
```

'On Calcul le taux de croissance incluant la volatilité ainsi que le choc aleatoire

' Calcul du growth rate avec un choc aleatoire moyen

```
Worksheets("Recapitulatif").Cells(21, 1) = "ExpectedGrowthRate"
```

```
Worksheets("Recapitulatif").Cells(21, 1).Interior.ColorIndex = 5
```

```
Worksheets("Recapitulatif").Cells(21, 1).Font.Bold = True
```

```
Worksheets("Recapitulatif").Cells(22, 1).Font.Bold = True
```

```
Worksheets("Recapitulatif").Cells(23, 1).Font.Bold = True
```

```
Worksheets("Recapitulatif").Cells(24, 1).Font.Bold = True
```

```
Worksheets("Recapitulatif").Cells(25, 1).Font.Bold = True
```



```
Worksheets("Recapitulatif").Cells(21, 1).Font.ColorIndex = 2
```

```
Worksheets("Recapitulatif").Cells(22, 1) = "Copper"
```

```
Worksheets("Recapitulatif").Cells(23, 1) = "Aluminium"
```

```
Worksheets("Recapitulatif").Cells(24, 1) = "TIN"
```

```
Worksheets("Recapitulatif").Cells(25, 1) = "HRC"
```

```
For i = 2 To 5
```

```
For h = 2 To Z + 1
```

```
Sheets("Recapitulatif").Cells(21, h) = " Month" & h - 1
```

```
Sheets("Recapitulatif").Cells(21, h).Interior.ColorIndex = 5
```

```
Sheets("Recapitulatif").Cells(21, h).Font.ColorIndex = 2
```

```
Sheets("Recapitulatif").Cells(21, h).Font.Bold = True
```

```
Sheets("Recapitulatif").Cells(i + 20, h) = Sheets("Recapitulatif").Cells(i + 10, h) +  
Sheets("Recapitulatif").Cells(i, h) * Sheets("Statistiques NbrAlea").Cells(i, h)
```

```
Sheets("Recapitulatif").Cells(i + 20, h).Interior.ColorIndex = 16
```

```
Sheets("Recapitulatif").Cells(i + 20, h).Font.ColorIndex = 2
```

```
Next h
```

```
Next i
```

```
' Calcul du growth rate avec un choc aleatoire pour les différents rate
```

```
For w = 1 To 4
```

```
For i = 2 To 5
```

```
For h = 2 To Z + 1
```

Worksheets("Recapitulatif").Cells(31, 1) = "GrowthRatePercentile " &  
 Sheets("Initialisation").Cells(8, 5)

Worksheets("Recapitulatif").Cells(41, 1) = "GrowthRatePercentile " &  
 Sheets("Initialisation").Cells(5, 5)

Worksheets("Recapitulatif").Cells(51, 1) = "GrowthRatePercentile " &  
 Sheets("Initialisation").Cells(5, 7)

Worksheets("Recapitulatif").Cells(61, 1) = "GrowthRatePercentile " &  
 Sheets("Initialisation").Cells(8, 7)

Worksheets("Recapitulatif").Cells(21 + 10 \* w, 1).Interior.ColorIndex = 5

Worksheets("Recapitulatif").Cells(21 + 10 \* w, 1).Font.Bold = True

Worksheets("Recapitulatif").Cells(22 + 10 \* w, 1).Font.Bold = True

Worksheets("Recapitulatif").Cells(23 + 10 \* w, 1).Font.Bold = True

Worksheets("Recapitulatif").Cells(24 + 10 \* w, 1).Font.Bold = True

Worksheets("Recapitulatif").Cells(25 + 10 \* w, 1).Font.Bold = True

Worksheets("Recapitulatif").Cells(21 + 10 \* w, 1).Font.ColorIndex = 2

Worksheets("Recapitulatif").Cells(22 + 10 \* w, 1) = "Copper"

Worksheets("Recapitulatif").Cells(23 + 10 \* w, 1) = "Aluminium"

Worksheets("Recapitulatif").Cells(24 + 10 \* w, 1) = "TIN"

Worksheets("Recapitulatif").Cells(25 + 10 \* w, 1) = "HRC"

Sheets("Recapitulatif").Cells(21 + 10 \* w, h) = " Month" & h - 1

Sheets("Recapitulatif").Cells(21 + 10 \* w, h).Interior.ColorIndex = 5

```
Sheets("Recapitulatif").Cells(21 + 10 * w, h).Font.ColorIndex = 2
```

```
Sheets("Recapitulatif").Cells(21 + 10 * w, h).Font.Bold = True
```

```
Sheets("Recapitulatif").Cells(i + 20 + 10 * w, h) = Sheets("Recapitulatif").Cells(i + 10, h)
+ Sheets("Recapitulatif").Cells(i, h) * Sheets("Statistiques NbrAlea").Cells(i + 10 * w, h)
```

```
Sheets("Recapitulatif").Cells(i + 20 + 10 * w, h).Interior.ColorIndex = 16
```

```
Sheets("Recapitulatif").Cells(i + 20 + 10 * w, h).Font.ColorIndex = 2
```

```
Next h
```

```
Next i
```

```
Next w
```

```
Columns("A:A").EntireColumn.AutoFit
```

```
'***** On effectue des prévisions*****'
```

```
ActiveWorkbook.Worksheets.Add.Move After:=Sheets(Sheets.Count)
```

```
Sheets(Sheets.Count).Name = "Forecast"
```

```
For w = 0 To 4
```

```
For i = 2 To 5
```

```
For h = 2 To Z + 1
```

Worksheets("Forecast").Cells(1, 1) = "Forecast Expected"

Worksheets("Forecast").Cells(11, 1) = "Forecast Percentile " &  
Sheets("Initialisation").Cells(8, 5)

Worksheets("Forecast").Cells(21, 1) = "Forecast Percentile " &  
Sheets("Initialisation").Cells(5, 5)

Worksheets("Forecast").Cells(31, 1) = "Forecast Percentile " &  
Sheets("Initialisation").Cells(5, 7)

Worksheets("Forecast").Cells(41, 1) = "Forecast Percentile " &  
Sheets("Initialisation").Cells(8, 7)

Worksheets("Forecast").Cells(1 + 10 \* w, 1).Interior.ColorIndex = 5

Worksheets("Forecast").Cells(1 + 10 \* w, 1).Font.Bold = True

Worksheets("Forecast").Cells(2 + 10 \* w, 1).Font.Bold = True

Worksheets("Forecast").Cells(3 + 10 \* w, 1).Font.Bold = True

Worksheets("Forecast").Cells(4 + 10 \* w, 1).Font.Bold = True

Worksheets("Forecast").Cells(5 + 10 \* w, 1).Font.Bold = True

Worksheets("Forecast").Cells(1 + 10 \* w, 1).Font.ColorIndex = 2

Worksheets("Forecast").Cells(2 + 10 \* w, 1) = "Copper"

Worksheets("Forecast").Cells(3 + 10 \* w, 1) = "Aluminium"

Worksheets("Forecast").Cells(4 + 10 \* w, 1) = "TIN"

Worksheets("Forecast").Cells(5 + 10 \* w, 1) = "HRC"

Sheets("Forecast").Cells(1 + 10 \* w, h) = " Month" & h - 1

```
Sheets("Forecast").Cells(1 + 10 * w, h).Interior.ColorIndex = 5
```

```
Sheets("Forecast").Cells(1 + 10 * w, h).Font.ColorIndex = 2
```

```
Sheets("Forecast").Cells(1 + 10 * w, h).Font.Bold = True
```

```
If h - 1 = 1 Then
```

```
    Sheets("Forecast").Cells(i + 10 * w, h) = Sheets("Initialisation").Cells(40 + i, 5) * (1  
+ Sheets("Recapitulatif").Cells(20 + i + 10 * w, h))
```

```
ElseIf h - 1 > 1 Then
```

```
    Sheets("Forecast").Cells(i + 10 * w, h) = Sheets("Forecast").Cells(i + 10 * w, h - 1) *  
(1 + Sheets("Recapitulatif").Cells(20 + i + 10 * w, h))
```

```
End If
```

```
Sheets("Forecast").Cells(i + 10 * w, h).Interior.ColorIndex = 16
```

```
Sheets("Forecast").Cells(i + 10 * w, h).Font.ColorIndex = 2
```

```
Next h
```

```
Next i
```

```
Next w
```

' On Calcule les moyennes mobiles pour Copper & Aluminium

'Insertion Dans Expected

Sheets("Forecast").Rows("3:3").Select

Selection.Insert Shift:=xlDown

Sheets("Forecast").Cells(3, 1) = "MovingAvg. " & Sheets("Forecast").Cells(2, 1)

Sheets("Forecast").Rows("5:5").Select

Selection.Insert Shift:=xlDown

Sheets("Forecast").Cells(5, 1) = "MovingAvg. " & Sheets("Forecast").Cells(4, 1)

'Insertion dans percentile 5%

Sheets("Forecast").Rows("15:15").Select

Selection.Insert Shift:=xlDown

Sheets("Forecast").Cells(15, 1) = "MovingAvg. " & Sheets("Forecast").Cells(2, 1)

Sheets("Forecast").Rows("17:17").Select

Selection.Insert Shift:=xlDown

Sheets("Forecast").Cells(17, 1) = "MovingAvg. " & Sheets("Forecast").Cells(4, 1)

'Insertion dans percentile 16%

Sheets("Forecast").Rows("27:27").Select

Selection.Insert Shift:=xlDown

Sheets("Forecast").Cells(27, 1) = "MovingAvg. " & Sheets("Forecast").Cells(2, 1)

Sheets("Forecast").Rows("29:29").Select

Selection.Insert Shift:=xlDown

```
Sheets("Forecast").Cells(29, 1) = "MovingAvg. " & Sheets("Forecast").Cells(4, 1)
```

```
'Insertion dans percentile 84%
```

```
Sheets("Forecast").Rows("39:39").Select
```

```
Selection.Insert Shift:=xlDown
```

```
Sheets("Forecast").Cells(39, 1) = "MovingAvg. " & Sheets("Forecast").Cells(2, 1)
```

```
Sheets("Forecast").Rows("41:41").Select
```

```
Selection.Insert Shift:=xlDown
```

```
Sheets("Forecast").Cells(41, 1) = "MovingAvg. " & Sheets("Forecast").Cells(4, 1)
```

```
'Insertion dans percentile 95%
```

```
Sheets("Forecast").Rows("51:51").Select
```

```
Selection.Insert Shift:=xlDown
```

```
Sheets("Forecast").Cells(51, 1) = "MovingAvg. " & Sheets("Forecast").Cells(2, 1)
```

```
Sheets("Forecast").Rows("53:53").Select
```

```
Selection.Insert Shift:=xlDown
```

```
Sheets("Forecast").Cells(53, 1) = "MovingAvg. " & Sheets("Forecast").Cells(4, 1)
```

```
For w = 0 To 4
```

```
    For h = 2 To Z + 1
```

```
        If h <= 7 Then
```

```
            Sheets("Forecast").Cells(3 + 12 * w, h) =
```

```
Application.WorksheetFunction.Average(Range(Sheets("Forecast").Cells(2 + 12 * w, 2),
```

```
Sheets("Forecast").Cells(2 + 12 * w, h)))
```

```

    Sheets("Forecast").Cells(5 + 12 * w, h) =
Application.WorksheetFunction.Average(Range(Sheets("Forecast").Cells(4 + 12 * w, 2),
Sheets("Forecast").Cells(4 + 12 * w, h)))

```

```

    ElseIf h > 7 Then

```

```

        Set OurRange = Range(Sheets("Forecast").Cells(2 + 12 * w, h - 6),
Sheets("Forecast").Cells(2 + 12 * w, h - 1))

```

```

        Set OurRange2 = Range(Sheets("Forecast").Cells(4 + 12 * w, h - 6),
Sheets("Forecast").Cells(4 + 12 * w, h - 1))

```

```

        Sheets("Forecast").Cells(3 + 12 * w, h) =
Application.WorksheetFunction.Average(OurRange)

```

```

        Sheets("Forecast").Cells(5 + 12 * w, h) =
Application.WorksheetFunction.Average(OurRange2)

```

```

    End If

```

```

    Next h

```

```

Next w

```

'On insere les prix actuels

```

Sheets("Forecast").Columns("B:B").Select

```

```

Selection.Insert Shift:=xlToRight

```

```

For w = 0 To 4

```

```

    For i = 2 To 7

```

```

        Sheets("Forecast").Cells(1 + 12 * w, 2) = "Prix Actuel"

```



Sheets("Forecast").Cells(1 + 12 \* w, 2).Interior.ColorIndex = 5

Sheets("Forecast").Cells(1 + 12 \* w, 2).Font.ColorIndex = 2

Sheets("Forecast").Cells(1 + 12 \* w, 2).Font.Bold = True

Sheets("Forecast").Cells(i + 12 \* w, 2).Interior.ColorIndex = 16

Sheets("Forecast").Cells(i + 12 \* w, 2).Font.ColorIndex = 2

Sheets("Forecast").Cells(i + 12 \* w, 2).Font.Bold = False

Sheets("Forecast").Cells(2 + 12 \* w, 2) = Sheets("Initialisation").Cells(42, 5)

Sheets("Forecast").Cells(3 + 12 \* w, 2) = Sheets("Initialisation").Cells(42, 5)

Sheets("Forecast").Cells(4 + 12 \* w, 2) = Sheets("Initialisation").Cells(43, 5)

Sheets("Forecast").Cells(5 + 12 \* w, 2) = Sheets("Initialisation").Cells(43, 5)

Sheets("Forecast").Cells(6 + 12 \* w, 2) = Sheets("Initialisation").Cells(44, 5)

Sheets("Forecast").Cells(7 + 12 \* w, 2) = Sheets("Initialisation").Cells(45, 5)

Next i

Next w

Columns("A:A").EntireColumn.AutoFit

\*\*\*On regroupe toutes les informations et on effectue une prévision de l'indice WPU10\*\*\*\*'

' On insere une feuille et on cree notre tableau de prevision

ActiveWorkbook.Worksheets.Add.Move After:=Sheets(Sheets.Count)

Sheets(Sheets.Count).Name = "Wpu10 Forecast"

For i = 2 To 4

For w = 0 To 4

For h = 2 To Z + 1

Worksheets("Wpu10 Forecast").Cells(1 + 12 \* w, 1) = Sheets("Forecast").Cells(1 + 12 \* w, 1)

Worksheets("Wpu10 Forecast").Cells(1 + 12 \* w, 1).Interior.ColorIndex = 5

Worksheets("Wpu10 Forecast").Cells(1 + 12 \* w, 1).Font.Bold = True

Worksheets("Wpu10 Forecast").Cells(1 + 12 \* w, 1).Font.ColorIndex = 2

Worksheets("Wpu10 Forecast").Cells(1 + 12 \* w, h) = " Month" & h - 1

Worksheets("Wpu10 Forecast").Cells(1 + 12 \* w, h).Interior.ColorIndex = 5

Worksheets("Wpu10 Forecast").Cells(1 + 12 \* w, h).Font.Bold = True

Worksheets("Wpu10 Forecast").Cells(1 + 12 \* w, h).Font.ColorIndex = 2

Worksheets("Wpu10 Forecast").Cells(1 + 12 \* w, 1).Font.Bold = True

Worksheets("Wpu10 Forecast").Cells(2 + 12 \* w, 1).Font.Bold = True

Worksheets("Wpu10 Forecast").Cells(3 + 12 \* w, 1).Font.Bold = True

Worksheets("Wpu10 Forecast").Cells(4 + 12 \* w, 1).Font.Bold = True

Worksheets("Wpu10 Forecast").Cells(5 + 12 \* w, 1).Font.Bold = True

Worksheets("Wpu10 Forecast").Cells(6 + 12 \* w, 1).Font.Bold = True

Worksheets("Wpu10 Forecast").Cells(7 + 12 \* w, 1).Font.Bold = True

Worksheets("Wpu10 Forecast").Cells(9 + 12 \* w, 1).Font.Bold = True

Worksheets("Wpu10 Forecast").Cells(2 + 12 \* w, 1) = "Copper"

Worksheets("Wpu10 Forecast").Cells(3 + 12 \* w, 1) = "Copper 6Mths MvgAvg"

Worksheets("Wpu10 Forecast").Cells(4 + 12 \* w, 1) = "Aluminium "

Worksheets("Wpu10 Forecast").Cells(5 + 12 \* w, 1) = "Aluminium 6Mths MvgAvg"

Worksheets("Wpu10 Forecast").Cells(6 + 12 \* w, 1) = "TIN"

Worksheets("Wpu10 Forecast").Cells(7 + 12 \* w, 1) = "HRC"

Worksheets("Wpu10 Forecast").Cells(9 + 12 \* w, 1) = "WPU10"

Worksheets("Wpu10 Forecast").Cells(2 + 12 \* w, h) = Sheets("Initialisation").Cells(70 + 2, 2) \* (Sheets("Forecast").Cells(2 + 12 \* w, h + 1) / Sheets("Forecast").Cells(2 + 12 \* w, h))

Worksheets("Wpu10 Forecast").Cells(3 + 12 \* w, h) = Sheets("Initialisation").Cells(70 + 3, 2) \* (Sheets("Forecast").Cells(3 + 12 \* w, h + 1) / Sheets("Forecast").Cells(3 + 12 \* w, h))

Worksheets("Wpu10 Forecast").Cells(4 + 12 \* w, h) = Sheets("Initialisation").Cells(70 + 4, 2) \* (Sheets("Forecast").Cells(4 + 12 \* w, h + 1) / Sheets("Forecast").Cells(4 + 12 \* w, h))

Worksheets("Wpu10 Forecast").Cells(5 + 12 \* w, h) = Sheets("Initialisation").Cells(70 + 5, 2) \* (Sheets("Forecast").Cells(5 + 12 \* w, h + 1) / Sheets("Forecast").Cells(5 + 12 \* w, h))

Worksheets("Wpu10 Forecast").Cells(6 + 12 \* w, h) = Sheets("Initialisation").Cells(70 + 6, 2) \* (Sheets("Forecast").Cells(6 + 12 \* w, h + 1) / Sheets("Forecast").Cells(6 + 12 \* w, h))

Worksheets("Wpu10 Forecast").Cells(7 + 12 \* w, h) = Sheets("Initialisation").Cells(70 + 7, 2) \* (Sheets("Forecast").Cells(7 + 12 \* w, h + 1) / Sheets("Forecast").Cells(7 + 12 \* w, h))

Next h

Next w

Next i

For k = 2 To Z + 1

```

    Sheets("Wpu10 Forecast").Cells(9, k).Value =
Application.WorksheetFunction.Sum(Sheets("Wpu10 Forecast").Range(Cells(2, k), Cells(8, k)))
+ Sheets("Initialisation").Cells(79, 2)

```

```

    Sheets("Wpu10 Forecast").Cells(9 + 12, k) =
Application.WorksheetFunction.Sum(Sheets("Wpu10 Forecast").Range(Cells(2 + 12, k), Cells(8
+ 12, k))) + Sheets("Initialisation").Cells(79, 2)

```

```

    Sheets("Wpu10 Forecast").Cells(9 + 24, k) =
Application.WorksheetFunction.Sum(Sheets("Wpu10 Forecast").Range(Cells(2 + 24, k), Cells(8
+ 24, k))) + Sheets("Initialisation").Cells(79, 2)

```

```

    Sheets("Wpu10 Forecast").Cells(9 + 36, k) =
Application.WorksheetFunction.Sum(Sheets("Wpu10 Forecast").Range(Cells(2 + 36, k), Cells(8
+ 36, k))) + Sheets("Initialisation").Cells(79, 2)

```

```

    Sheets("Wpu10 Forecast").Cells(9 + 48, k) =
Application.WorksheetFunction.Sum(Sheets("Wpu10 Forecast").Range(Cells(2 + 48, k), Cells(8
+ 48, k))) + Sheets("Initialisation").Cells(79, 2)

```

```

    Sheets("Wpu10 Forecast").Cells(9 + 60, k) =
Application.WorksheetFunction.Sum(Sheets("Wpu10 Forecast").Range(Cells(2 + 60, k), Cells(8
+ 60, k))) + Sheets("Initialisation").Cells(79, 2)

```

```

    Sheets("Wpu10 Forecast").Cells(9 + 72, k) =
Application.WorksheetFunction.Sum(Sheets("Wpu10 Forecast").Range(Cells(2 + 72, k), Cells(8
+ 72, k))) + Sheets("Initialisation").Cells(79, 2)Next k

```

```

Columns("A:A").EntireColumn.AutoFit

```

```

End Sub

```